Income Inequality, House Prices, and Housing Regulations *

Chao Liu[†] Edmund Y. Lou[‡] Wei Xiang[§]

April 13, 2024

Abstract

This paper studies the relationship between income inequality and house prices in the United States. We exploit the initial income distribution across 90 occupation-income percentile groups and the national income growth of these groups to construct instruments for income inequality measures, effectively addressing reverse causality concerns. Using county-level data from 1990 to 2017, we find that a one standard deviation increase in the Gini coefficient leads to a 26% increase in house prices. We propose a supply-side channel to explain the observed higher prices and fewer housing stocks in more unequal areas. Consistent with this mechanism, our analysis shows that a one standard deviation increase in the Gini coefficient results in an increase in the 2018 Wharton Residential Land Use Regulation Index by 0.35 standard deviations, leading to 14% fewer housing units, 58% fewer building permits over the next decade, and a 2 percentage point decrease in homeownership rate. Our findings highlight the importance of income inequality in shaping housing market dynamics through its impact on housing regulations and the consequences for housing affordability.

*We thank Kaiji Chen, Anthony DeFusco, Daniel Greenwald, Adam Guren, Aaron Hedlund, Zhikun Lu, Gregor Matvos, John Mondragon, Charles G. Nathanson, and Paola Sapienza, as well as seminar participants at Northwestern Kellogg.

†Liu: Kellogg School of Management, Northwestern University. Email: chao.liu1@kellogg.northwestern.edu.

‡Lou: Department of Economics, Northwestern University. Email: edmund.lou@u.northwestern.edu.

§Xiang: Department of Economics, Yale University. Email: wei.xiang@yale.edu.

Income inequality has risen dramatically in the United States over the past few decades (Autor, Katz and Kearney, 2008; Hoffmann, Lee and Lemieux, 2020; Saez and Zucman, 2020). The same period also witnessed a significant increase in house prices, particularly during the lead-up to the Great Recession (Knoll, Schularick and Steger, 2017). Figure 1 illustrates the striking co-movements of income inequality, as measured by the Gini coefficient, and the house price index in the U.S. from 1990 to 2018. The Gini coefficient increased from 0.43 in 1990 to 0.49 in 2018, while the house price index reached nearly 2.7 times its 1990 level in 2018, albeit with a notable boom-bust episode around the financial crisis. The close relationship between these variables raises a fundamental question: Is there a causal relationship between income inequality and house prices?

Despite a growing body of literature exploring this relationship, the causal impact of income inequality on house prices remains ambiguous, as the direction and magnitude of the effect depend on the underlying mechanisms and the data employed by researchers. On one hand, rising income inequality may lead to higher house prices if it increases the demand for housing among wealthy households or prompts them to invest more in housing as a store of wealth (Zhang, 2016). On the other hand, income inequality could lead to lower house prices if it dampens the housing demand among low- and middle-income households (Kösem, 2023). The inconclusive evidence in the existing literature underscores the need for a rigorous investigation of the causal link between income inequality and house prices.

Understanding the interplay between income inequality and housing market outcomes is not only of academic interest but also has far-reaching policy implications. If income inequality indeed causes higher house prices, it may exacerbate wealth disparities, given that homeownership is a major source of wealth accumulation for many households. Rising house prices could make it more difficult for low- and middle-income households to access the benefits of homeownership (Haurin, Parcel and Haurin, 2002; Goodman and Mayer, 2018; Sodini et al., 2023). Moreover, housing wealth has significant implications for the well-being of households and the stability of the financial system through its impact on household consumption and borrowing behaviors (Gan, 2010; Mian and Sufi, 2011; Bhutta and Keys, 2016). Therefore, understanding this relationship is essential for designing effective policies that promote inclusive growth and maintain financial stability.

This paper attempts to estimate the causal effect of income inequality on house prices and explore the mechanisms through which this relationship operates. To accompolish this, we compile a comprehensive county-level panel data set spanning from 1990 to 2017, which combines information on income inequality, house prices, housing supply, and regulatory environments from various sources. In particular, we uncover a valuable yet underutilized data source, the Census Historical Income Tables, which provides comprehensive income distribution metrics at the county level for the years 1990 and 2000. This unique data set allows us to construct various measures of income inequality, such as the Gini coefficient, income percentile ratios, and top income shares.

In exploring the relationship between housing prices and income inequality, our study tackles a significant identification challenge: reverse causality, where income inequality might both influence and be influenced by house prices. To disentangle this bidirectional relationship, we first investigate the impact of housing prices on income inequality using an instrumental variable (IV) approach. We employ the land unavailability index developed by Lutz and Sand (2023) as an instrument for housing prices. This index captures the geographic constraints on housing supply across counties, providing a plausibly exogenous source of variation in housing prices. Our IV estimates suggest that rising housing prices lead to a reduction in the Gini coefficient, likely because low-income households are forced to relocate to more affordable areas. This finding has important implications for interpreting the ordinary least squares (OLS) estimates of the effect of income inequality on housing prices, as it indicates that the OLS estimates are likely to be biased downward due to reverse causality.

To address this reverse causality and identify the causal effect of income inequality on housing prices, we develop a novel IV for income inequality. Our instrument is constructed by exploiting the initial income distribution across 90 occupation-income groups in 1980 and the national growth patterns of these groups over time. By construction, this IV is uncorrelated with local migration patterns, mitigating concerns about reverse causality. Moreover, the predicted income distribution allows us to construct other income inequality indicators besides the Gini coefficient. The validity of this Bartik-style IV relies on the exogeneity of initial shares, requiring that they do not directly affect future house prices through channels other than future income inequality (Goldsmith-Pinkham, Sorkin and Swift, 2020). We argue that some initial shares, such as the shares of military workers in certain income bins, are likely to be exogenous as they were determined by national policies in the 1980s rather than current local economic conditions. Furthermore, we develop two tests to demonstrate that all initial shares in our setting contribute relatively equally to the overall identification, suggesting that no single share drives the results. The likely exogeneity of some shares, combined with the equal contribution of all shares, strengthens the case for the overall exogeneity of our instrument.

Using this IV strategy, we find that a one standard deviation increase in the Gini coefficient leads to a 26% increase in house prices. This effect is particularly strong in counties with inelastic housing supply, suggesting that the responsiveness of housing supply plays a crucial role in shaping the relationship between income inequality and house prices. Consistent with the negative effect of house prices on income inequality, the IV estimates are larger than the OLS estimates. Our main results are robust to alternative measures of income inequality and sample selections.

The second question this paper aims to answer is through which channels income inequality increases housing prices. While prior literature has focused on potential demand-side channels (Gyourko, Mayer and Sinai, 2013; Zhang, 2016; Kösem, 2023), we provide suggestive evidence highlighting the necessity of supply-side mechanisms. Using our IV approach, we show that greater income inequality is associated with higher prices but fewer housing units across U.S. counties, implying that demand-side mechanisms alone cannot fully explain how income inequality shapes housing market equilibrium.

We propose a supply-side channel operating through housing regulations, hypothesizing that higher income inequality leads to stricter housing regulations, which constrain housing supply and drive up housing prices. The case of Marc Andreessen, a prominent billionaire, illustrates this channel. He publicly advocates for the construction of additional housing to address the critical issue of housing affordability (Andreessen, 2020):

> *"It's Time to Build...crazily skyrocketing housing prices in places like San Francisco, making it nearly impossible for regular people to move in and take the jobs of the future."*

Despite this call to action, Andreessen (2022) has asked for the removal of all multifamily zoning projects near his residence, citing concerns that such developments would:

> *"MASSIVELY decrease our home values, the quality of life of ourselves and our neighbors and IMMENSELY increase the noise pollution and traffic."*

The apparent contradiction in Andreessen's stance exemplifies how wealthy individuals in unequal areas may oppose housing development to protect their property values and quality of life, aligning with our hypothesis that higher income inequality can lead to stricter housing regulations.

We provide empirical evidence to support our proposed supply-side mechanism. First, we examine a recent policy change in California, the Senate Bill 35 (SB 35), which aimed to expedite the development of affordable housing units. We find that more unequal districts were more likely to oppose SB 35, indicating that income inequality could influence the political support for restrictive housing policies. Second, to study this supply-side channel more systematically, we use the Wharton Residential Land Use Regulatory Index (WRLURI), which measures the stringency of local housing regulations across the U.S (Gyourko, Saiz and Summers, 2008; Gyourko, Hartley and Krimmel, 2021). Our analysis finds that a one standard deviation increase in the Gini coefficient is associated with a 0.35 standard deviation increase in the 2018 WRLURI, suggesting that higher income inequality leads to more restrictive housing regulations. Finally, we estimate the direct effect of income inequality on housing supply, as measured by the issuance of building permits. Our results suggest that a one standard deviation increase in the Gini coefficient results in a 58% decline in building permits over the next decade. As a consequence of higher prices and reduced supply of new housing, we find that the same change in the Gini coefficient leads to a decrease in homeownership rates by approximately 2 percentage points.

This paper is related to several bodies of literature. First, this paper contributes to the strand of literature examining the relationship between income inequality and housing markets. Gyourko, Mayer and Sinai (2013) demonstrate that an increasing number of high-income households in the U.S. intensifies competition for limited land resources, driving up housing prices in superstar cities. Similarly, Zhang (2016) uses data from China to show that investment motive among the wealthy plays a key role in house price formation. In contrast, Määttänen and Terviö (2014) argue that house prices are primarily determined by non-high-income households, suggesting that increased income inequality may lower house prices. Kösem (2023) also argue that a rise in income inequality can lead to a decline in housing prices via a borrower risk composition channel. Our paper makes two contributions to this debate. First, we use an IV strategy to estimate the causal relationship between income inequality and housing prices in both directions. Second, by uncovering a supply-side mechanism, we contribute to a better understanding of the complex relationship between income inequality, housing regulations, and housing affordability. Our findings have important policy implications, as they suggest that addressing income inequality and reforming housing regulations may be crucial for promoting housing affordability and reducing wealth inequality.

Second, our paper adds to the studies on the causes and consequences of housing regulations. While a large body of research has investigated the effect of land use regulations (Glaeser and Gyourko, 2003; Bunten, 2017; Glaeser and Gyourko, 2018), studies on the causes of these regulations are relatively sparse. Glaeser and Ward (2009) find that while historical housing density strongly predicts current minimum lot sizes in Greater Boston, the causes of other housing regulations are difficult to explain with observable characteristics. Parkhomenko (2023) develops a quantitative spatial equilibrium model where local regulation is determined endogenously by voting. Landowners in productive cities with attractive amenities vote for strict regulation. Our paper complements these studies by providing empirical evidence showing that higher levels of income inequality are associated with more stringent housing regulations, shedding light on an important factor influencing the adoption of restrictive housing policies.

Lastly, our paper contributes to the extensive literature on the socioeconomic effects of inequality. Income inequality has been linked to various adverse outcomes. For example, Fajnzylber, Lederman and Loayza (2002) and Enamorado et al. (2016) document a positive relationship between income inequality and violent crime. Pickett and Wilkinson (2015) conduct a literature review within an epidemiological causal framework and find that inequality leads to poorer health and lower life expectancy. Furthermore, Galor and Zeira (1993) argue that inequality can lead to underinvestment in human capital among lower-income households. However, the effects of income inequality remain contested. Welch (1999) argues that inequality is necessary for incentivizing effort and innovation and allocating talent to its most productive uses. Boustan et al. (2013) find that growing income inequality is associated with an expansion in public expenditures on a wide range of services. Our paper focuses on the housing market, which has the potential to shape economic inequality in multiple dimensions. Changes in the housing market can be a key driver for increased inequality in income after housing expenditure, leading to a divergence in consumption and savings patterns across income groups (Etheridge, 2019; Dustmann, Fitzenberger and Zimmermann, 2022). Methodologically, building on the work of Boustan et al. (2013), we develop new instruments for income inequality that can be applied in other studies. We discuss the plausibility of our Bartik-style IV in line with the recent applied econometrics literature (Goldsmith-Pinkham, Sorkin and Swift, 2020).

The remainder of this paper is organized as follows. Section I outlines our data sources and presents stylized facts about income inequality and housing markets in the U.S. Section II details our empirical strategy and our main results about the effect of income inequality on housing prices. Section III introduces a novel supply-side mechanism and provides direct empirical evidence on the impact of inequality on housing regulations and supply. Section IV concludes.

I Data and Aggregate Facts

A Data

We compile a decennial county-level panel using various data sources outlined below, offering detailed information on local income inequality and housing market dynamics from 1990 to 2017.

Income inequality: To measure income inequality in 1990 and 2000, we use the Census Historical Income Tables.[1] This data set provides comprehensive income distribution metrics, such as quintile income thresholds, average income per quintile, and quintile percentage shares, alongside the commonly used Gini coefficients. Furthermore, we collect the same set of inequality metrics from the American Community Survey (ACS) 5-year Estimates. Specifically, we use data from the 2008-2012 and 2015-2019 periods for 2010 and 2017, respectively.

Table 1 presents summary statistics for four income inequality indicators that we mainly use in this paper: the Gini coefficient, the income share of the top 20 percent, and the ratios of mean income between the highest and third quintiles, and between the third and lowest quintiles. In 1990, the average Gini coefficient was 0.415, which rose to 0.445 in 2017. The observed increases across all four measures indicate a general growth in income inequality throughout our study period. Notably, the widening income disparity was more pronounced between the wealthiest and the middle-income groups than between the middle-income and the poorest groups. Figure 2 displays the county-level distributions of these income inequality measures in 1990 and 2017.

[1]The data are available at `https://www.census.gov/data/tables/time-series/dec/historical-income-counties.html`. To our knowledge, this data set has been underutilized in previous literature. Researchers typically use state-level inequality measures from Frank (2014) (e.g., Kösem (2023)) or calculate county-level Gini coefficients from grouped income data (e.g., Boustan et al. (2013)).

Except for the ratio between the mean income levels of the third and first quintiles, the distributions of the other three measures significantly shifted rightward from 1990 to 2017.

House prices: The Federal Housing Finance Agency (FHFA) House Price Index serves as our principal data source for house prices. This index tracks average price fluctuations through repeat sales or refinancings of identical single-family homes whose mortgages have been purchased or securitized by Fannie Mae or Freddie Mac. This paper employs both national and county-level indices. Additionally, we use the median housing value from the census and ACS as a supplementary data source.

Housing supply: We quantify housing supply through the number of building permits issued by local authorities for new privately-owned residential construction projects. The statistics come from the Building Permits Survey (BPS), which compiles reports from local building permit officials. To smooth out any year-to-year idiosyncrasies in housing supply, we aggregate the number of building permits issued over subsequent ten-year intervals. However, this approach means that we lack data for this variable for the year 2017. As shown in Table 1, there is a discernible decline in the issuance of building permits between 1990 and 2017.

Housing regulations: The Wharton Residential Land Use Regulatory Index (WRLURI) captures regulatory restrictions in the housing market. Based on a survey conducted in 2005, the index sheds light on various aspects of local land use control environments, including the general characteristics of the regulatory process, statutory limits on development, density restrictions, open space requirements, infrastructure cost sharing and approval delay (Gyourko, Saiz and Summers, 2008). The final product, WRLURI2006, is derived from the principal component analysis of 11 survey-based measures, representing the strictness of local zoning regulations. This index is normalized to have a zero mean and a unit standard deviation, with higher values indicating stricter regulatory environments. For analytical purposes, we aggregate the raw data at the community level to the county level.

In a subsequent iteration, Gyourko, Hartley and Krimmel (2021) applied the original methodology to create the WRLURI2018, reflecting contemporary regulatory landscapes. Our study merges the WRLURI2006 with data from the year 2000 and the WRLURI2018 with data from the year 2017. Due to the fact that only 890 communities participated in both the initial and follow-up surveys, we analyze the WRLURI2006 and WRLURI2018 indices independently rather than combining them into a longitudinal data set.

Supplemental data: The county-level ACS files are our sources for county-year level characteristics, including demographic features, housing stocks, and homeownership rates. For a comprehensive view of the mortgage market, we turn to the Home Mortgage Disclosure Act (HMDA) database, which covers the near universe of mortgage activities in the United States. After excluding refinancing mortgages, we use the mortgage origination volume as a proxy for housing demand. The proportion of applications for non-owner-occupied properties serves as an indicator of investment demand in the housing market.[2] As shown in Table 1, the volume of mortgage originations mirrors the national boom and bust housing cycle. About 12.6% of mortgage applications were for loans on non-owner-occupied properties in our sample.

B Aggregate Facts

Having described our data, we next present several stylized facts on income inequality, housing values, housing stocks, regulatory landscapes, housing supply, and homeownership rates. We use binned scatter plots to highlight notable correlations among these variables in the pooled sample, setting the stage for a more rigorous investigation of their causal relationships in the following sections. To rule out the income effect while presenting patterns in the raw data, we only control for the real mean household income (adjusted to 1990 dollars) and total population when the dependent variable measures an aggregate quantity.

Panel A of Figure 3 illustrates a strongly positive relationship between the Gini coefficient and the logarithm of real median housing values (adjusted to 1990 dollars). The displayed line represents a linear regression of housing values on the Gini coefficients and mean household income across the entire data set.[3] Figure B1 extends this analysis to four distinct time points in our dataset. Except for in 2000, we observe a persistent positive correlation, with coefficients of linear fits significant at the 1% level.

Next, we examine the relationship between income inequality and housing stocks, measured by the logarithm of owner-occupied housing units. Given the strong correlation between housing stocks and population size, we control for total population in our analysis. As shown in Panel B of Figure 3, after accounting for income and population levels, counties with higher levels of income inequality typically have fewer

[2] Gao, Sockin and Xiong (2020) use this variable to measure speculation in the housing market.

[3] We also report the coefficient for the Gini coefficient and its standard error in parentheses. The coefficient is presented solely to demonstrate a significant correlation; it is not intended to quantify the magnitude of the relationship.

owner-occupied housing units. Figure B2 confirms this pattern across four distinct time periods within our dataset.

Residential land use regulations may have an impact on housing stocks over time by constraining the supply of new housing units, so we next explore the association between income inequality and housing regulations. The left figure in Panel C of Figure 3 suggests that counties with higher Gini coefficients in 2000 tended to have higher Wharton Regulation Index values (i.e., stricter housing regulations). Likewise, the right figure indicates a positive correlation between income inequality in 2010 and the stringency of residential land use regulation in 2018.

Panel D of Figure 3 demonstrates a negative correlation between income inequality and housing supply, which is measured by the number of building permits issued over the next ten years scaled by current housing stocks. Cross-sectional binned scatter plots could provide a clearer visualization of this relationship. As depicted in Figure B3, counties with higher levels of inequality in 2000 and 2010 initiated fewer residential construction projects in the subsequent decade. This pattern is consistent with the previous finding, as these counties also had a higher Wharton Regulation Index during the 2000s and 2010s. Despite the absence of regulatory data from the 1990s, a negative correlation between income inequality and housing supply persists.

Finally, given the above stylized facts, it is not surprising that income inequality has an adverse effect on housing affordability. Panel E of Figure 3 shows a strongly negative relationship between the Gini coefficients and homeownership rates in our pooled sample. Furthermore, as shown in Figure B4, we consistently observe lower homeownership rates in counties with higher Gini coefficients throughout the last four decades.

II Income Inequality and House Prices

A Empirical Strategy

Motivated by the pooled and repeated cross-sectional evidence in the last section, we next use county-level panel regressions with fixed effects to explore the relationship between income inequality and house prices. Specifically, the relationship of interest is described as the following equation:

$$log(HPI_{ct}) = \alpha + \beta Inequality_{ct} + \Gamma X_{ct} + \delta_c + \delta_{st} + \varepsilon_{ct}, \quad (1)$$

where c, s, and t denote county, state, and year respectively. The dependent variable, $log(HPI_{ct})$, is the log transformation of the FHFA Housing Price Index for county c in year t. X_{ct} represents a set of time-varying local characteristics, including mean household income, total population, unemployment rate, the proportion of minorities, and the share of population with post-secondary education. Our preferred specification controls for mean income, which captures the idea of a mean preserving spread in income distribution. As a robustness check, we also examine models that control for median income, which in part determines median voters' preferences for local housing regulation policies.[4]. In Equation (1), we add county fixed effects δ_i so that the variation in the dependent variable comes from the change in income inequality of the same county. We also include state-year fixed effects, δ_{st}, to account for common shocks to the housing market happening at the state-year level, such as shifts in tax legislation or banking regulations. Our coefficient of interest, β, captures the effect of local income inequality on the housing price index.

A.1 Threat to Identification

Equation (1) alone does not suffice to establish a causal relationship between income inequality and local house prices. Income distribution could affect local housing prices through various mechanisms, such as the speculative activities of wealthy households or the preferences of local voters on residential land use. Conversely, the change in housing prices might induce shifts in the local income distribution. For example, rising house prices could force lower-income individuals to relocate to more affordable areas, potentially reducing the local income inequality. This migration pattern could lead to a downward bias in the OLS estimates of the effect of income inequality on house prices. Alternatively, an increase in house prices could attract wealthy households to the area, thereby widening the local income distribution. In this scenario, the bias direction of the OLS estimates becomes ambiguous.

To address this threat to identification more precisely, we examine the possibility of reverse causality–namely, how house prices affect income inequality. This relationship is captured by the following equation:

$$Inequality_{ct} = \alpha_r + \beta_r log(HPI_{ct}) + \Gamma_r X_{ct} + \delta_c + \delta_t + \varepsilon_{ct}. \tag{2}$$

[4] Boustan et al. (2013) control for median household income when studying the effect of income inequality on taxation and public expenditures. Similarly, Enamorado et al. (2016) use this specification for their main results to estimate the effect of income inequality on violent crime in Mexico.

Since our focus is the effect of income inequality on housing price, we need an IV for housing price to estimate Equation (2). Housing supply elasticity has been widely used in previous literature for this purpose. A popular candidate is the Saiz elasticity IV (Saiz, 2010), which combines land unavailable due to geographic building restrictions with local housing market regulations.[5] However, the Saiz data set only contains a small number of cities, making it less applicable to our study due to the geographical mismatch and limited data coverage.[6] Therefore, we turn to the novel data set of Land Unavailability (LU) index developed by Lutz and Sand (2023). Using accurate satellite imagery and modern machine learning techniques, this data set captures the geographic determinants of housing supply for nearly all U.S. counties. Given the time-invariant nature of this variable, we interact it (i.e., $Unavailability_c$) with the national housing price index ($log(HPI_t)$) as our IV for local housing prices ($log(HPI_{ct})$).

Table B1 reports the estimates of the first-stage regression. We regress local housing price index on our IV without controls in column 1, then introduce county controls in column 2, and finally include fixed effects in column 3. In all specifications, the first-stage coefficients are statistically significant at the 1% level. This pattern suggests that counties with inelastic land supply tend to experience exaggerated responses to the national housing market cycles.

We next estimate Equation (2) using both OLS and 2SLS methods. Table 2 shows that house prices have a negative impact on income inequality, with OLS results in odd columns and IV results in even columns. The F-statistic from the first stage exceeds 52, which alleviates weak instrument concerns. The first two columns demonstrates that rising house prices are correlated with reduced poverty rates. Our IV estimate in column 2 suggests that a 100% increase in housing prices could lead to a decrease in poverty rates by 6.2 percentage points. Columns 3 and 4 indicate that high housing price areas do not see a full replacement of departing low-income households with wealthier newcomers at the top end of the income distribution. According to column 4, the income share of top 20 percent would decrease by 0.034 if house prices increase by 100%. The last two columns use a comprehensive measure of income inequality, the Gini coefficient, as the dependent variable. As shown in column 6, we find that an increase in housing price by 100% could lead to a decline in the Gini coefficient

[5] Recent applications include Mian and Sufi (2009, 2011); Chaney, Sraer and Thesmar (2012); Mian and Sufi (2014); Aladangady (2017).

[6] Other critiques of the Saiz IV include Davidoff (2016)'s concern about potential correlation with local housing demand and Guren et al. (2021)'s observation of its limited first-stage predictive power.

by roughly 0.027. Notably, our IV estimates are all smaller than the OLS estimates. If income inequality has a positive causal impact on house prices, such upward bias is anticipated. Otherwise, the effect of income inequality on housing prices must be substantially negative. The estimate of β in Equation (1) should be smaller than -37, which means that a 0.036 increase in the Gini coefficient (one standard deviation in our sample) must reduce local housing price by over 133%. These results suggest a positive effect of income inequality on house prices and a downward bias in the OLS estimates from Equation (1).

A.2 Instrumental Variable for Income Inequality

To address concerns about reverse causality, we develop an IV that is associated with a county's Gini coefficient but is otherwise uncorrelated with house prices. The construction of our IV is similar to that of the shift-share IVs, which have been extensively used in applied labor economics studies (Bartik, 1991; Blanchard et al., 1992). Building on and refining the method introduced by Boustan et al. (2013), we predict future income distribution of a county based on its initial income distribution (initial share component) and national patterns of income growth (shift component).[7] We then calculate the Gini coefficient for this predicted distribution and use it as an instrument for the actual Gini coefficient. This IV approach also helps mitigate measurement error in local Gini coefficients, which may introduce an attenuation bias in the OLS estimates.

The construction of our IV proceeds as follows.[8] First, we determine the initial share component of the instrument. We start from the county-level employment of 15 occupations in 1980, a decade earlier than our data for analysis.[9] For example, suppose that in 1980, a percent of people in county c of state s had an occupation o. Next, we exploit the national income distribution for each occupation in 1980 to further divide the state-level occupation employment into 6 income percentile bins (i.e., 0-10, 10-30, 30-70, 70-90, 90-98, 98-100). For instance, let us say that b percent of people with occupation o in state s earned an income above the 98th percentile of the national income distribution for this occupation (i.e., 98-100th percentile group). By combining these two steps, we obtain the initial shares of 90 (15 occupations $\times$ 6

[7]Appendix A provides a detailed comparison of our IV construction with that of Boustan et al. (2013).

[8]We provide a concrete example of the procedure in Appendix A.

[9]Table A1 presents the 15 occupations and their corresponding occupation codes (variable *OCC1990*) from the IPUMS U.S. Census 1980 dataset.

income percentile bins) occupation-income groups for each county, denoted as $f_{ic,1980}$, where i represents an occupation-income group. In our example, we would assume that $a \times b$ percent of households in county c belong to the occupation-income group i, which is characterized by occupation o and income above the 98th percentile of the national income distribution of this occupation. It is important to note that, unlike traditional Bartik instruments, we incorporate state-level information into the local initial shares, as the distribution of income percentile bins for each occupation is assumed to be the same for all counties within a given state.

Second, we derive the shift component of our instrument. Using micro-level census data, we follow a leave-one-out strategy for each state to calculate the national median income of each occupation-income group. This approach addresses the finite sample bias that may arise from using own-observation information. Specifically, for each state s and each occupation-income group i, we calculate the national income level at the midpoint of group i, excluding data from state s. For example, consider the occupation-income group i (i.e., the 98-100 percentile income group of occupation o) in year t (later than 1980). For counties in state s, we calculate the 99th percentile of the national income distribution for occupation o in year t, excluding data from state s. This process yields a set of national median incomes for each occupation-income group i and year t, denoted as $\tilde{w}_{it,-s}$, where t takes the value of 1990, 2000, 2010, or 2017, and $-s$ indicates that the data from state s is excluded.

Given the initial shares ($f_{ic,1980}$) and national shifts ($\tilde{w}_{it,-s}$), we can predict the income distribution for future periods by assuming that the income level of each occupation-income group grows over time according to the observed national changes, while the shares of occupation-income groups remain fixed at their 1980 levels. In our example, we assume that in county c of state s in year t, $a \times b$ percent of households belonging to occupation-income group i have an income level of $\tilde{w}_{it,-s}$. Using the predicted income distribution, we can calculate the predicted Gini coefficient as

$$Gini_{ct}^{IV} = \frac{1}{2\bar{\bar{\tilde{W}}}_{ct}} \sum_{i=1}^{K} \sum_{j=1}^{K} f_{ic,1980} f_{jc,1980} \left| \tilde{w}_{it,-s} - \tilde{w}_{jt,-s} \right|, \tag{3}$$

where $\bar{\bar{\tilde{W}}}_{ct}$ is the predicted mean income in county c in year t, which can be calculated as $\sum_{i=1}^{K} f_{ic,1980} \times \tilde{w}_{it,-s}$. By construction, this IV is uncorrelated with local migration patterns, mitigating concerns about reverse causality. Moreover, the predicted income distribution allows us to construct other income inequality indicators besides the Gini coefficient. In the following sections, we present the IV estimation results and conduct several robustness checks to assess the validity of our instrument.

B Empirical Results

Our 2SLS regressions are as follows:

$$Gini_{ct} = \alpha_f + \phi_f Gini_{ct}^{IV} + \Gamma_f X_{ct} + \delta_c + \delta_{st} + \varepsilon_{ct}, \tag{4}$$

$$log(HPI_{ct}) = \alpha_{iv} + \beta_{iv} \widehat{Gini}_{ct} + \Gamma_{iv} X_{ct} + \delta_c + \delta_{st} + \varepsilon_{ct}. \tag{5}$$

We find a robust and positive relationship between the predicted and actual Gini coefficients, supporting the validity of our IV approach. Figure 4 presents the first-stage correlation in both levels and changes, demonstrating that our IV captures the underlying dynamics in local income inequality well. Column 1 of Table 3 shows that the coefficient for the first-stage relationship is statistically significant at the 1% level, providing further empirical support for this relationship. The F-statistic on the predicted Gini coefficient is 35.57 in the full sample, substantially surpassing the conventional threshold for a strong IV (Stock and Yogo, 2005).

Using the predicted Gini coefficients as an instrument, we find that income inequality has a positive impact on local house prices. Column 2 of Table 3 reports our main IV result. The coefficient indicates that a one standard deviation increase in the Gini coefficient (about 0.036) leads to a 26% rise in housing prices. From 1990 to 2017, the average Gini coefficient increased by 0.03, while the housing price index surged by roughly 238.4%. Thus, our IV estimate suggests that escalating income inequality could explain nearly 8% of the housing price increase over this period. In columns 3 and 4, we split the sample by the proportion of land unavailable for construction at the county level, based on the index provided by Lutz and Sand (2023). The effect of income inequality on housing prices is more pronounced in counties with relatively inelastic land supply, as indicated by the larger coefficient magnitudes and higher statistical significance in column 3. This finding aligns with the intuition that constrained housing supply limits the ability to adjust housing prices downward in these markets.

Our IV design is crucial for estimating the causal effect of income inequality. The last column of Table 3 reports the OLS estimate of Equation (1). The resulting estimate of β is negative, consistent with the findings of Kösem (2023).[10] Table B3

[10] Table B2 reports our OLS estimates following the specification in Kösem (2023). Although

demonstrates that this downward bias is not specific to the Gini coefficient. Other indicators of income inequality also suffer from the potential reverse causality problem. Interestingly, if we control for median income rather than mean income, as shown in Table B4, the OLS estimates become positive and statistically significant. The instability of the OLS results provides suggestive evidence of the misspecification in the OLS model and highlights the need for an IV strategy. The substantial downward bias of the OLS estimates relative to the IV estimates suggests that housing prices have a negative impact on income inequality. We have provided empirical evidence supporting this reverse relationship in Table 2. By employing an IV strategy, we can identify the causal effect of income inequality on housing prices, mitigating the confounding influence of reverse causality and potential measurement error in income inequality metrics.

To check the robustness of our main results presented above, we use a variety of alternative specifications. Since we predict the entire income distribution for later years, we can construct IVs for other inequality measures, such as the income share of the top 20 percent and the ratio between the mean income of the highest and third quintiles. Figure B5 shows a strong positive correlation between the predicted and actual measures. We apply these two additional IVs for income inequality to supplement our main analysis using the Gini coefficient. As shown in Table B5, our preferred IV estimate implies that a 1 percentage point increase in the income share of the top 20 percent would lead to an approximately 10% higher housing price. In Table B6, we use the ratio between the mean income of the highest and third quintiles to measure inequality and find a significantly positive effect of inequality on housing prices. Consistent with our main findings, we observe a larger impact in counties with inelastic land supply when using these two alternative inequality indicators.

We also show that our main results are robust to different sample selection criteria. To mitigate the impact of the Great Recession, We drop observations from 2010 and repeat our analysis. Table B7 shows that our findings are robust to this exercise, with results very similar to our baseline estimates. Additionally, when we estimate Equation (5) using a balanced panel, the results reported in Table B8 are consistent with our baseline findings. The balanced panel consists of relatively large counties with better data coverage, indicating that our findings are not driven by county size. As another robustness check, we exclude counties that are not part of metropolitan

the granularity of our data and the sample period differ from her study, we find very similar OLS estimates.

areas from our sample. As Table B9 shows, the coefficients maintain their expected signs and remain statistically significant, confirming that our findings are not driven by urban/rural differences. Taken together, these robustness checks provide further support for the positive causal relationship between income inequality and housing prices. The consistency of our results across different inequality measures, model specifications, and sample selection criteria underscores the reliability of our findings.

C Discussion on the Instrumental Variable

Goldsmith-Pinkham, Sorkin and Swift (2020) prove that using a Bartik instrument is equivalent to using initial shares as multiple instruments, so the exogeneity condition should be interpreted in terms of these shares. In our context, the initial (1980) shares are local shares of 15 occupations, each with 6 income bins (i.e., $f_{ic,1980}$). Building on Rotemberg (1983), the authors decompose the Bartik estimator into a weighted sum of the just-identified IV estimators that use each initial share as as a separate IV. They suggest that researchers should focus on shares with larger Rotemberg weights, which indicate the sensitivity of the overall estimate to misspecification (i.e., endogeneity) in each share. However, as evident from Equation (3), our instrument does not take the canonical form of a Bartik IV, which is an inner product of initial shares and national shifts. Consequently, we cannot directly calculate these Rotemberg weights and identify the most influential shares.

Instead of attempting to identify which shares receive more weight and are worth testing for the identifying assumptions, we aim to demonstrate that all initial shares in our setting contribute relatively equally to the overall identification. We believe that some shares are more likely to be exogenous. For instance, it seems implausible that the share of military workers in the 10th to 30th income percentiles in 1980 could directly affect future house prices, as military employment is largely determined by national defense policies and is unlikely to be correlated with local housing market conditions. If this plausibly exogenous share contributes similarly to identification as the other shares, it would support the validity of our approach. To formalize this idea, we develop two tests to show that our IV results are not driven by specific initial shares.

First, we intentionally drop a set of potentially influential occupation-income groups when constructing the predicted Gini coefficient. We consider four different sets of excluded groups: (1) the 10 groups with the highest income, (2) the 10 groups with the lowest income, (3) the 10 groups with the largest employment, and

(4) a combination of the 5 groups with the highest income and the 5 groups with the lowest income. Table 4 presents the 2SLS results from these four alternative instrument constructions. Remarkably, the first-stage F-statistics remain quite large in all cases, indicating that the instrument remains strong even after excluding these potentially influential groups. Moreover, we continue to find a significantly positive impact of the Gini coefficient on house prices across all specifications. Crucially, the magnitude of the coefficients is similar to that of our baseline results, which use all 90 occupation-income groups to construct the instrument.

Second, we randomly select 80 groups from all 90 occupation-income groups to calculate the predicted Gini coefficient, and then we use this new IV to run our 2SLS regressions. We repeat this exercise 500 times. Panel A of Figure 5 plots the first-stage F-statistics of the 500 iterations. Remarkably, only one of them has an F-statistic smaller than ten, indicating that the instrument remains strong even when using a subset of all occupation-income groups. Panel B of Figure 5 plots the IV coefficients of the 500 iterations using the partially simulated IVs. The coefficients fluctuate around our baseline results, and about 95% of them are statistically significant at the 5% level. This exercise demonstrates that our IV estimates are not sensitive to the specific set of occupation-income groups used to construct the instrument.

The number of occupation-income groups we choose in constructing the instrument determines the tradeoff between the first-stage strength and the plausibility of the exclusion restriction. As we increase the number of groups used, the predictive power of our IV improves, making it more likely to meet the relevance requirement. However, using more groups also increases the risk of including influential groups that may violate the exclusion restriction. Figures B6 and B7 plot the results of the randomization tests when using 70 and 60 groups, respectively. Together with Figures 5, these figures illustrate the tradeoff.

When we use only 60 groups to construct the instrument, 40 iterations have a first-stage F-statistic falling below the conventional threshold of ten (Panel A of Figure B7). Nevertheless, the majority of the coefficients remain statistically significant and are centered around our baseline estimate (Panel B of Figure B7). Figure B6 presents an intermediate case, where we use 70 groups to construct the instrument. There are 16 first-stage F-statistics smaller than ten. The distribution of the IV coefficients in Panel B of Figure B6 is more tightly centered around our baseline estimate compared to Figure B7, but it exhibits slightly more variation than Figure 5.

Taken together, both types of tests suggest that there are no specific initial shares disproportionately influencing our results. The robustness of our findings across these exercises reinforces the validity of our empirical strategy. In the following sections, we will continue to use this IV approach to examine the effect of income inequality on other important variables in local housing markets.

III A Supply-Side Mechanism

Our findings in the previous section establish a positive causal relationship between income inequality and housing prices, but the underlying mechanisms remain unclear. To shed light on this question, we proceed in three steps in this section. First, we demonstrate the necessity of supply-side mechanisms to reconcile the observed equilibrium in housing markets. Second, we provide direct empirical evidence on the supply-side channel, which are understudied in previous literature. Finally, we discuss the implications of our findings for housing affordability.

A Necessity of Supply-Side Mechanisms

The striking co-movements of household debt, income inequality, and house prices before the 2008 financial crisis have motivated a prevalent view that the rise in income inequality might have caused some of the increase in household leverage and thus led to higher house prices (Rajan, 2011). However, recent studies have offered evidence not entirely consistent with this argument. For example, Coibion et al. (2020) find that low-income households in high-inequality regions accumulated less debt relative to their income than low-income households in lower inequality regions. Kösem (2023) uses a panel of U.S. states from 1992 to 2015 to show that a rise in income inequality led to a higher share of borrowers opting for risky mortgages and lower mortgage debt.

Using our IV strategy, we directly investigate the causal effect of income inequality on mortgage origination volume, which serves as a proxy for housing demand. Columns 2 and 3 of Table B10 show that income inequality has a negative effect on mortgage origination, consistent with the findings of Kösem (2023). Our IV estimate implies that an increase in the Gini coefficient by 0.01 is associated with an 83% decrease in mortgage origination volume. This result suggests that rising income inequality may not necessarily lead to increased housing demand, hinting at the importance of supply-side factors in explaining the positive effect of income inequality on housing prices.

It is important to acknowledge that mortgage origination may not fully capture housing demand due to the existence of all cash buyers. In fact, it is possible that income inequality has a positive impact on housing demand. For instance, Zhang (2016) argues that a rise in inequality causes aggregate housing demand and prices to increase because wealthy households' demand is more responsive than that of the poor. To test this hypothesis, we examine the effect of income inequality on housing investment in columns 4 and 5 of Table B10. We use the share of mortgage applications for non-owner-occupied properties as a measure of investment motives. Our findings indicate that income inequality has a positive impact on housing investment. As shown in column 5, the IV estimate suggests that a 0.01 increase in the Gini coefficient is associated with a 6 percentage point increase in the share of investment applications.

If income inequality drives up wealthy people's housing demand, can this demand-side channel alone explain higher housing prices in more unequal regions? To better understand the implications of the above empirical findings, let us consider a standard market equilibrium framework, where the quantity and price of housing are determined by the intersection of supply and demand curves. If income inequality were to elevate house prices solely through demand-side mechanisms, we would expect to see a higher number of housing units in areas with greater income inequality. However, our empirical evidence reveals that higher income inequality is correlated with a lower number of housing units, suggesting that supply-side factors must be at play.

Table 5 reports our IV estimates of the effect of income inequality on housing stocks. While the OLS estimate in column 1 suggests a positive correlation between the Gini coefficient and the number of housing units, this is likely due to reverse causality, as housing stocks can also impact income inequality. If an area has sufficient housing units to accommodate households, it may have a higher Gini coefficient since low-income households do not have to move. Table B11 provides direct evidence supporting this argument. Using land unavailability as the instrument for the change in housing stocks from 1990 to 2017, column 3 of this table shows that a 10% increase in housing stocks will lead to an increase in the Gini coefficient by roughly 0.02.

Once we account for this reverse causality, as shown in column 2 of Table 5, income inequality actually has a negative impact on housing stocks. The IV estimate suggests that if the Gini coefficient increases by one standard deviation, there will be approximately 14% fewer housing units and 13% owner-occupied units. Housing units per capita will decrease by 0.06. This evidence underscores the importance of considering supply-side mechanisms when analyzing the impact of income inequality on house prices.

Previous literature has mainly focused on the demand-side effect of income inequality on housing markets, such as reduced mortgage origination among low-income households and increased housing investment among the wealthy. In this part, our results imply that regardless of the direction of the demand-side effect, income inequality must have an impact on the supply side to reconcile the fact that higher prices and fewer units are observed in more unequal areas.

B Effect on Housing Regulations and Supply

To shed light on the supply-side channel through which income inequality affects house prices, we begin by examining a recent policy change in California as an illustrative example. In 2017, the California State Legislature passed Senate Bill 35 (SB 35), which aimed to addressing the state's acute housing shortage. The bill introduces a streamlined approval process for housing development projects in cities that have not met state-mandated Regional Housing Needs Allocation (RHNA) goals. Figure B8 plots the voting results of state senate districts in California, with districts that voted "Yea" (or in favor) represented in darker blue. "Yea" votes were concentrated primarily in the coastal regions of California, particularly in the San Francisco Bay Area, Los Angeles, and San Diego metropolitan areas. By reducing the barriers to construction, SB 35 seeks to expedite the development of affordable housing units across California. This bill is part of a broader suite of housing-related laws enacted to encourage the construction of new housing stock and increase accessibility and affordability for residents in California.

To investigate whether income inequality is correlated with local support for this housing reform, we conduct a simple OLS analysis. As shown in Table B12, more unequal districts are more likely to vote against SB 35. The result is robust to different measures of income inequality: the Gini coefficient, the income share of top 20 percent, and the income share of top 5 percent. This finding provides anecdotal evidence that income inequality may influence house prices through its impact on housing regulation and housing supply.

To study this housing regulation channel more systematically, we use the Wharton Residential Land Use Regulation Index (WRLURI), a comprehensive measure of the stringency of local land use regulations across the United States. Panels A and B of

Table 6 present our results for WRLURI2006 and WRLURI2018, respectively. In the first two columns of each panel, the outcome variable is the Local Political Pressure Index (LPPI), which quantifies the degree to which various actors, such as local councils and local community groups, are involved in the local residential development process. Although a higher degree of public and local official involvement could arise from a desire to reduce the regulatory burden, it almost certainly reflects a relatively high level of existing restrictiveness (Gyourko, Hartley and Krimmel, 2021). The 2006 index is standardized to have a mean of 0 and a standard deviation of 1, while the 2018 index retains its original scale, ranging from 3 to 15. Our IV estimates imply that a one standard deviation (approximately 0.03) increase in the Gini coefficient would lead to an increase in LPPI by roughly 0.64 standard deviations in 2006 and 0.40 standard deviations in 2018.

Columns 3 and 4 focus on the Supply Restrictions Index (SRI), which reflects the extent to which there are explicit annual caps on the supply of new housing. This index is the sum of the number of limits on building permits, construction, or number of dwellings and units, with values ranging from a low of 0 to a high of 6 in both the 2006 and 2018 versions. As shown in column 4, our IV estimates indicate that a one standard deviation increase in the Gini coefficient would lead to an increase in SRI by approximately 0.34 standard deviations in 2006 and 0.27 standard deviations in 2018.

The dependent variable in the last two columns is the Wharton Residential Land Use Regulatory Index (WRLURI), which is the first factor from a variety of subindexes that gauge different components of the underlying regulatory environment.[11] This comprehensive index is standardized so that the sample mean is 0 and its standard deviation is 1, with lower (higher) values of the index reflecting a less (more) restrictive regulatory regime. Based on the IV estimates in column 6, we find that a one standard deviation increase in the Gini coefficient would lead to an increase in WRLURI by roughly 0.57 standard deviations in 2006 and 0.35 standard deviations in 2018.

Overall, our findings provide robust evidence that higher levels of income inequality are associated with more restrictive housing regulations, as measured by various indexes capturing different aspects of the regulatory environment. The results in two panels imply that the impact remains significant from 2006 to 2018, despite a slight decrease in magnitude. As shown in Table 6, the IV estimates are consistently larger than their OLS counterparts. This downward bias in the OLS estimates can be ex-

[11] The 2006 version consists of 11 subindexes, while the 2018 version includes 12 subindexes.

plained by the same reverse causality logic presented in Table 3. Stringent housing regulations, similar to high housing prices, may lead to a lower Gini coefficient as low-income households are compelled to relocate to more affordable areas.

To further strengthen our findings on the supply-side mechanisms, we examine the direct effect of income inequality on the issuance of building permits, which serves as a leading indicator measure of new housing construction. By investigating this relationship, we aim to assess whether the stringency of land use regulations, influenced by income inequality, translates into actual changes in housing supply. Table 7 reports our results. The dependent variable in the first two columns is the log of building permits issued in the following ten years. Our IV estimate in column 2 shows that a one standard deviation increase in the Gini coefficient would lead to a 58% decline in building permits issued over the next decade. The upward bias of the OLS estimate (column 1) is due to the reverse causality, where more housing supply could increase income inequality, a pattern consistent with our discussion in Table 5. In columns 3 and 4, we confirm the negative effect of inequality on housing supply, using per capita building permits issued in the following decade as the dependent variable. In columns 5 and 6, we focus on per capita building permits issued for projects with more than five units. This type of residential projects is more likely to be affordable housing projects, making it a suitable measure of housing supply for low- and middle-income households. As shown in column 6, a one standard deviation increase in the Gini coefficient is associated with a decrease in per capita multifamily building permits by 0.51 standard deviations. Our findings on building permits complement the previous results on land use regulations and provide a comprehensive picture of the supply-side channel linking income inequality to housing market outcomes.

C Implications on Housing Affordability

Having provided empirical evidence that higher income inequality leads to higher housing prices, reduced housing stocks, and decreased housing supply, it is reasonable to anticipate a lower homeownership rate in more unequal areas. To investigate the consequences of income inequality and constrained housing supply on housing affordability, we examine the effect of income inequality on homeownership rates.

Table 8 presents the results of our analysis. The F-statistics for the excluded instruments in the IV specifications are well above the conventional threshold, indicating the strength of our instruments. The OLS estimate in column 1 suggests that the Gini coefficient is negatively associated with homeownership rate. However, as discussed earlier, the OLS estimate may be biased due to endogeneity issues. Our preferred IV estimate in column 2 shows that a one standard deviation increase in the Gini coefficient leads to a decrease in homeownership rates by roughly 1.9 percentage points. Columns 3 and 4 present the IV estimates for counties with inelastic and elastic housing supply, respectively. We find that the negative effect of income inequality on homeownership rates is more pronounced in counties with inelastic housing supply. Specifically, a one standard deviation increase in the Gini coefficient reduces homeownership rates by 6.1 percentage points in inelastic counties, while the effect is statistically insignificant in elastic counties. This finding is consistent with our previous results, which demonstrates that the impact of income inequality on housing markets is more substantial in areas with constrained housing supply.

To further confirm the negative effect of income inequality on housing affordability, we explore the relationship between the Gini coefficient and the price-income ratio in Table B13. The ratio is calculated as the median housing value divided by the median household income, with a larger value indicating less affordability. The IV estimate in column 2 suggests that a one standard deviation increase in the Gini coefficient would result in an increase in the price-income ratio by 0.93 standard deviations. The evidence reported in columns 3 and 4 reinforces our conclusion that the impact of income inequality on housing markets is exacerbated in areas with limited housing supply elasticity.

IV Conclusion

This paper estimates the causal relationship between income inequality and housing prices by introducing a novel Bartik-style instrument for the Gini coefficient. This instrument is calculated based on a predicted county-level income distribution using the 1980 income distribution and national patterns of income growth. Our results indicate that increased income inequality leads to reduced housing units and higher housing prices, which cannot be explained solely via a demand-side mechanism.

We provide suggestive evidence on the supply-side channel by investigating the voting results on Senate Bill 35 in California. We document that districts with higher income inequality tend to vote against the law. Using our instrument for the Gini coefficient, we further find that rising income inequality is associated with more restrictive housing regulations, reduced supply of new housing, and lower homeownership rates. Therefore, a comprehensive understanding of why top-income households prefer more stringent housing regulations is crucial for designing policies that aim to expand housing supply or promote housing affordability.

Moving forward, we plan to develop a quantitative model wherein households across various income percentiles exhibit varying responses to the negative externalities stemming from housing density. Notably, high-income households tend to lobby for stricter regulations on housing supply. Armed with this model, we aim to assess the impact of different distributional policies on housing prices.

References

Aladangady, Aditya. 2017. "Housing Wealth and Consumption: Evidence from Geographically Linked Microdata." *American Economic Review* 107(11):3415–3446.

Andreessen, Marc. 2020. "It's Time to Build.". https://a16z.com/its-time-to-build.

Andreessen, Marc. 2022. "IMMENSELY AGAINST multifamily development!". https://www.econlib.org/immensely-confusing/.

Autor, David H., Lawrence F. Katz and Melissa S. Kearney. 2008. "Trends in U.S. Wage Inequality: Revising the Revisionists." *Review of Economics and Statistics* 90:300–323.

Bartik, Timothy J. 1991. "Who Benefits From State and Local Economic Development Policies?" *WE Upjohn Institute for Employment Research* .

Bhutta, Neil and Benjamin J Keys. 2016. "Interest Rates and Equity Extraction during the Housing Boom." *American Economic Review* 106(7):1742–1774.

Blanchard, Olivier Jean, Lawrence F. Katz, Robert E. Hall and Barry Eichengreen. 1992. "Regional Evolutions." *Brookings Papers on Economic Activity* 1992(1):1–75.

Boustan, Leah, Fernando Ferreira, Hernan Winkler and Eric M Zolt. 2013. "The Effect of Rising Income Inequality on Taxation and Public Expenditures: Evidence from US Municipalities and School Districts, 1970–2000." *Review of Economics and Statistics* 95(4):1291–1302.

Bunten, Devin. 2017. "Is the Rent Too High? Aggregate Implications of Local Land-Use Regulation." *Available at SSRN 2980048* .

Chaney, Thomas, David Sraer and David Thesmar. 2012. "The Collateral Channel: How Real Estate Shocks Affect Corporate Investment." *American Economic Review* 102(6):2381–2409.

Coibion, Olivier, Yuriy Gorodnichenko, Marianna Kudlyak and John Mondragon. 2020. "Greater Inequality and Household Borrowing: New Evidence From Household Data." *Journal of the European Economic Association* 18(6):2922–2971.

Davidoff, Thomas. 2016. "Supply Constraints Are Not Valid Instrumental Variables for Home Prices Because They Are Correlated With Many Demand Factors." *Critical Finance Review* 5(2):177–206.

Dustmann, Christian, Bernd Fitzenberger and Markus Zimmermann. 2022. "Housing Expenditure and Income Inequality." *The Economic Journal* 132(645):1709–1736.

Enamorado, Ted, Luis F López-Calva, Carlos Rodríguez-Castelán and Hernán Winkler. 2016. "Income Inequality and Violent Crime: Evidence from Mexico's Drug War." *Journal of Development Economics* 120:128–143.

Etheridge, Ben. 2019. "House Prices and Consumption Inequality." *International Economic Review* 60(4):1781–1822.

Fajnzylber, Pablo, Daniel Lederman and Norman Loayza. 2002. "Inequality and Violent Crime." *Journal of Law and Economics* 45(1):1–39.

Frank, Mark. 2014. "A New State-Level Panel of Annual Inequality Measures Over the Period 1916-2005." *Journal of Business Strategies* 31(1):241–263.

Galor, Oded and Joseph Zeira. 1993. "Income Distribution and Macroeconomics." *The Review of Economic Studies* 60(1):35–52.

Gan, Jie. 2010. "Housing Wealth and Consumption Growth: Evidence From a Large Panel of Households." *The Review of Financial Studies* 23(6):2229–2267.

Gao, Zhenyu, Michael Sockin and Wei Xiong. 2020. "Economic Consequences of Housing Speculation." *Review of Financial Studies* 33(11):5248–5287.

Glaeser, Edward and Joseph Gyourko. 2018. "The Economic Implications of Housing Supply." *Journal of Economic Perspectives* 32(1):3–30.

Glaeser, Edward L and Bryce A Ward. 2009. "The Causes and Consequences of Land Use Regulation: Evidence from Greater Boston." *Journal of Urban Economics* 65(3):265–278.

Glaeser, Edward L. and Joseph Gyourko. 2003. "The Impact of Building Restrictions on Housing Affordability." *Federal Reserve Bank of New York Economic Policy Review* 9:21–39.

Goldsmith-Pinkham, Paul, Isaac Sorkin and Henry Swift. 2020. "Bartik Instruments: What, When, Why, and How." *American Economic Review* 110(8):2586–2624.

Goodman, Laurie S and Christopher Mayer. 2018. "Homeownership and the American Dream." *Journal of Economic Perspectives* 32(1):31–58.

Guren, Adam M, Alisdair McKay, Emi Nakamura and Jón Steinsson. 2021. "Housing Wealth Effects: The Long View." *Review of Economic Studies* 88(2):669–707.

Gyourko, Joseph, Albert Saiz and Anita Summers. 2008. "A New Measure of the Local Regulatory Environment for Housing Markets: The Wharton Residential Land Use Regulatory Index." *Urban Studies* 45(3):693–729.

Gyourko, Joseph, Christopher Mayer and Todd Sinai. 2013. "Superstar Cities." *American Economic Journal: Economic Policy* 5(4):167–199.

Gyourko, Joseph, Jonathan S Hartley and Jacob Krimmel. 2021. "The Local Residential Land Use Regulatory Environment Across US Housing Markets: Evidence from a New Wharton Index." *Journal of Urban Economics* 124:103337.

Haurin, Donald R, Toby L Parcel and R Jean Haurin. 2002. "Does Homeownership Affect Child Outcomes?" *Real Estate Economics* 30(4):635–666.

Hoffmann, Florian, David S. Lee and Thomas Lemieux. 2020. "Growing Income Inequality in the United States and Other Advanced Economies." *Journal of Economic Perspectives* 34:52–78.

Knoll, Katharina, Moritz Schularick and Thomas Steger. 2017. "No Price Like Home: Global House Prices, 1870-2012." *American Economic Review* 107:331–353.

Kösem, Sevim. 2023. "Income Inequality, Mortgage Debt and House Prices." *Bank of England Staff Working Paper No. 921* .

Lutz, Chandler and Ben Sand. 2023. "Highly Disaggregated Land Unavailability." *Available at SSRN 3478900* .

Määttänen, Niku and Marko Terviö. 2014. "Income Distribution and Housing Prices: An Assignment Model Approach." *Journal of Economic Theory* 151:381–410.

Mian, Atif and Amir Sufi. 2009. "The Consequences of Mortgage Credit Expansion: Evidence from the US Mortgage Default Crisis." *Quarterly Journal of Economics* 124(4):1449–1496.

Mian, Atif and Amir Sufi. 2011. "House Prices, Home Equity–Based Borrowing, and the US Household Leverage Crisis." *American Economic Review* 101(5):2132–2156.

Mian, Atif and Amir Sufi. 2014. "What Explains the 2007–2009 Drop in Employment?" *Econometrica* 82(6):2197–2223.

Parkhomenko, Andrii. 2023. "Local Causes and Aggregate Implications of Land Use Regulation." *Journal of Urban Economics* 138:103605.

Pickett, Kate E and Richard G Wilkinson. 2015. "Income Inequality and Health: A Causal Review." *Social science & medicine* 128:316–326.

Rajan, Raghuram G. 2011. *Fault Lines: How Hidden Fractures Still Threaten the World Economy.* Princeton University Press.

Rotemberg, Julio. 1983. "Instrument Variable Estimation of Misspecified Models.".

Saez, Emmanuel and Gabriel Zucman. 2020. "The Rise of Income and Wealth Inequality in America: Evidence from Distributional Macroeconomic Accounts." *Journal of Economic Perspectives* 34:3–26.

Saiz, Albert. 2010. "The Geographic Determinants of Housing Supply." *Quarterly Journal of Economics* 125(3):1253–1296.

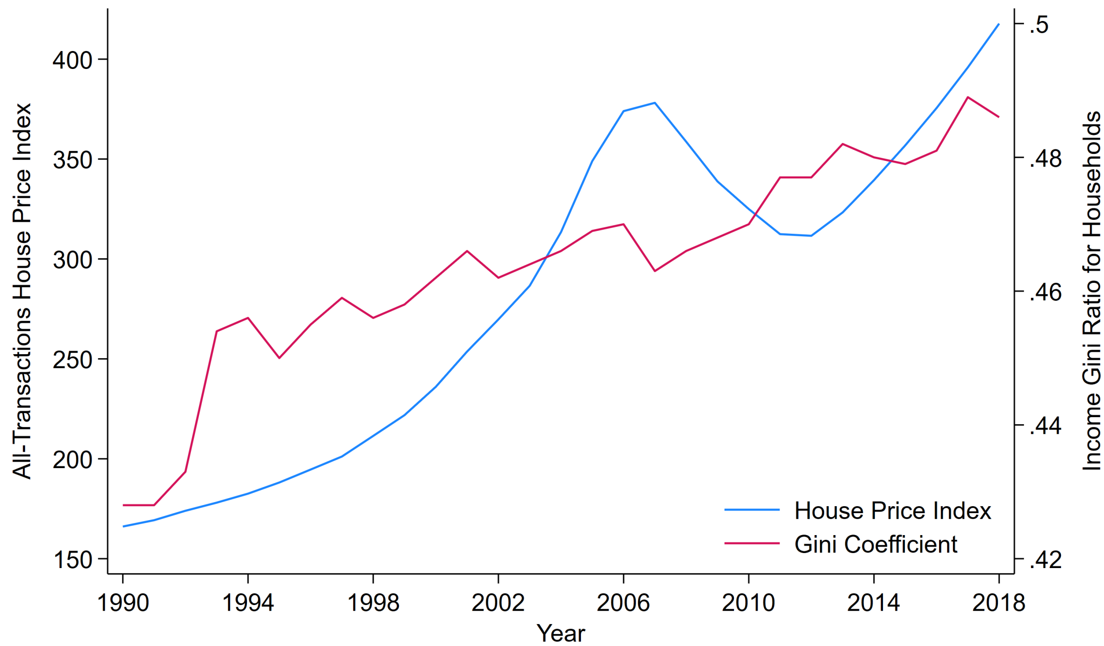
Sodini, Paolo, Stijn Van Nieuwerburgh, Roine Vestman and Ulf von Lilienfeld-Toal. 2023. "Identifying the Benefits from Homeownership: A Swedish Experiment." *American Economic Review* 113(12):3173–3212.

Stock, James H. and Motohiro Yogo. 2005. Testing for Weak Instruments in Linear IV Regression. In *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg*, ed. Donald W. K. Andrews and James H. Stock. New York: Cambridge University Press pp. 80–108.

Welch, Finis. 1999. "In Defense of Inequality." *American Economic Review* 89(2):1–17.

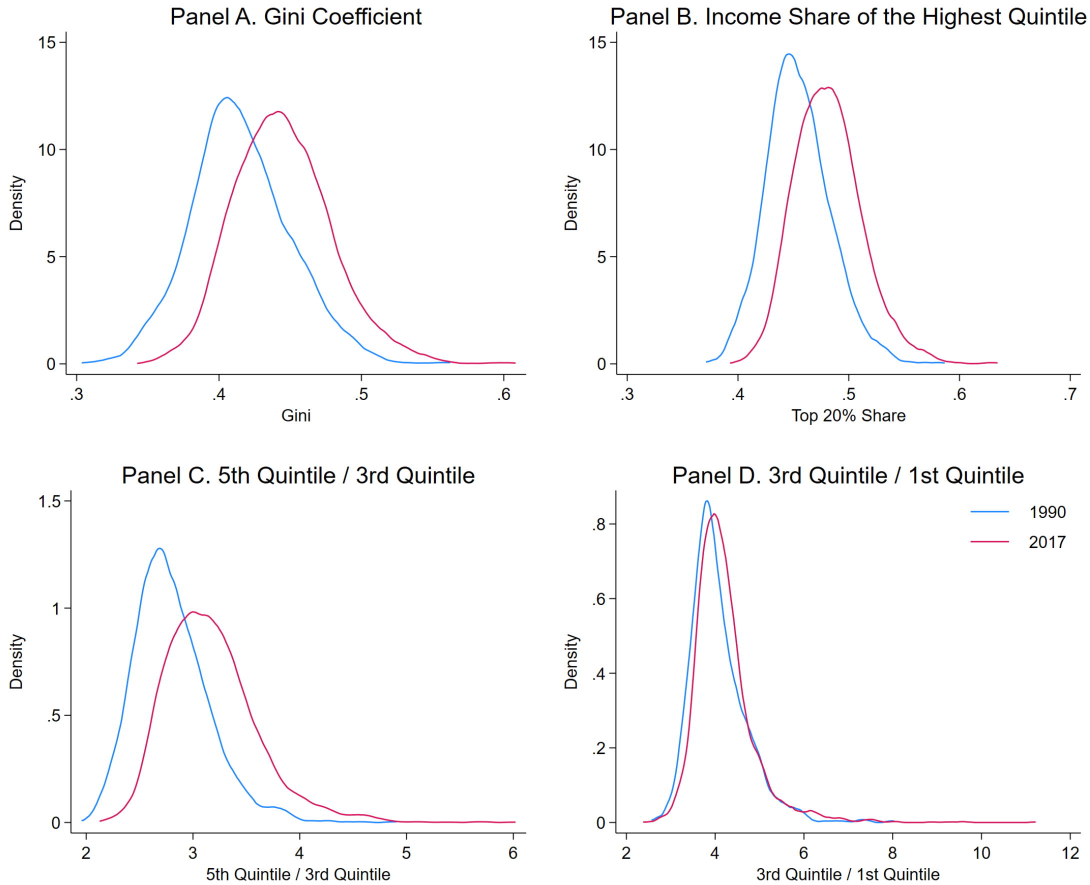
Zhang, Fudong. 2016. "Inequality and House Prices." Mimeo, University of Michigan.

Figure 1. Inequality and House Price Index in the U.S.



Notes: This figure plots the Gini coefficient and the house price index in the U.S. from 1990 to 2018. The Gini coefficient is obtained from the U.S. Census Bureau, while the house price index is sourced from the Federal Housing Finance Agency (FHFA). The index is normalized to 100 in the first quarter of 1980.

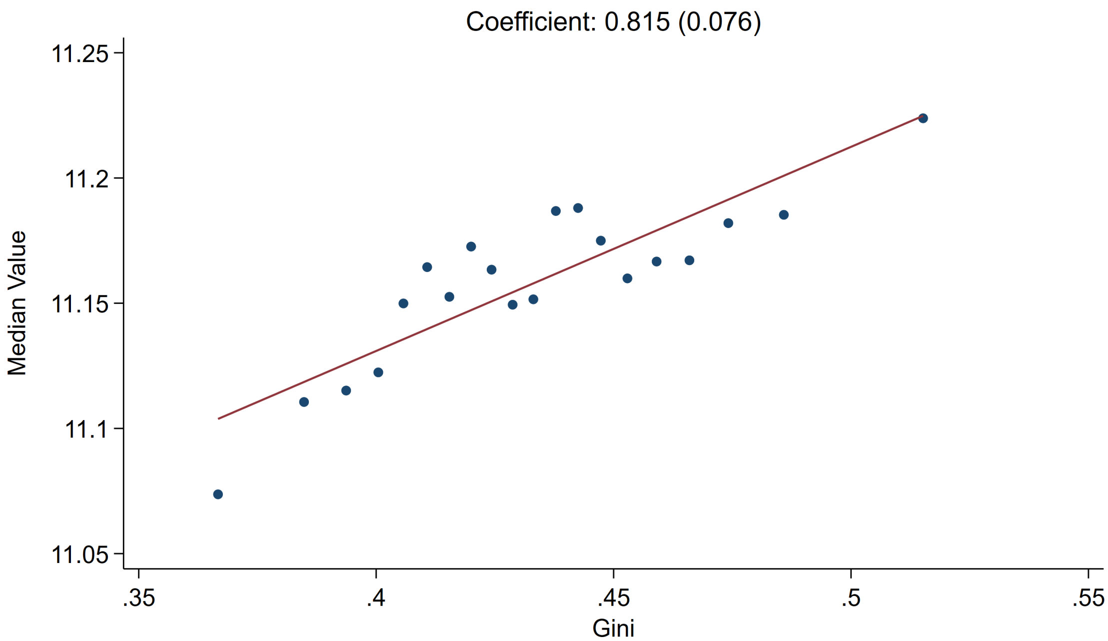
Figure 2. Distribution of Income Inequality in the U.S.



Notes: This figure plots the county-level distribution of four measures of income inequality for the years 1990 and 2017: the Gini coefficient (panel A), the income share of the top 20 percent (panel B), the ratio between the mean income of the highest and third quintiles (panel C), and the the ratio between the mean income of the third and lowest quintiles (panel D).

Figure 3. Correlation Between the Gini Coefficient and Other Variables

(a) Median Housing Value



(b) Owner Occupied Units

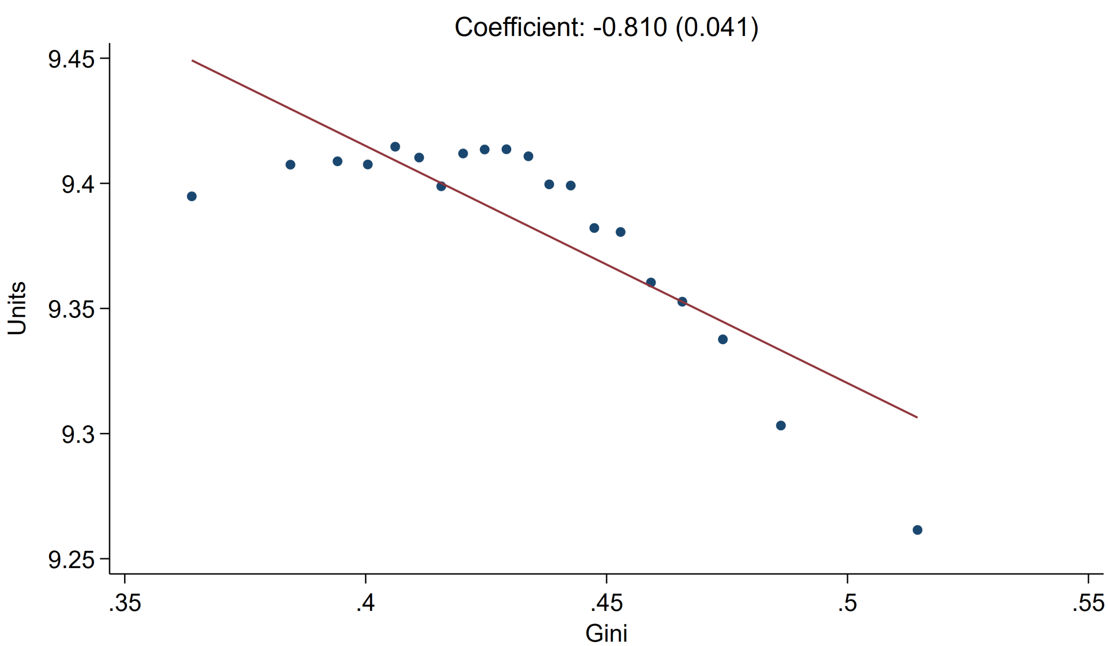
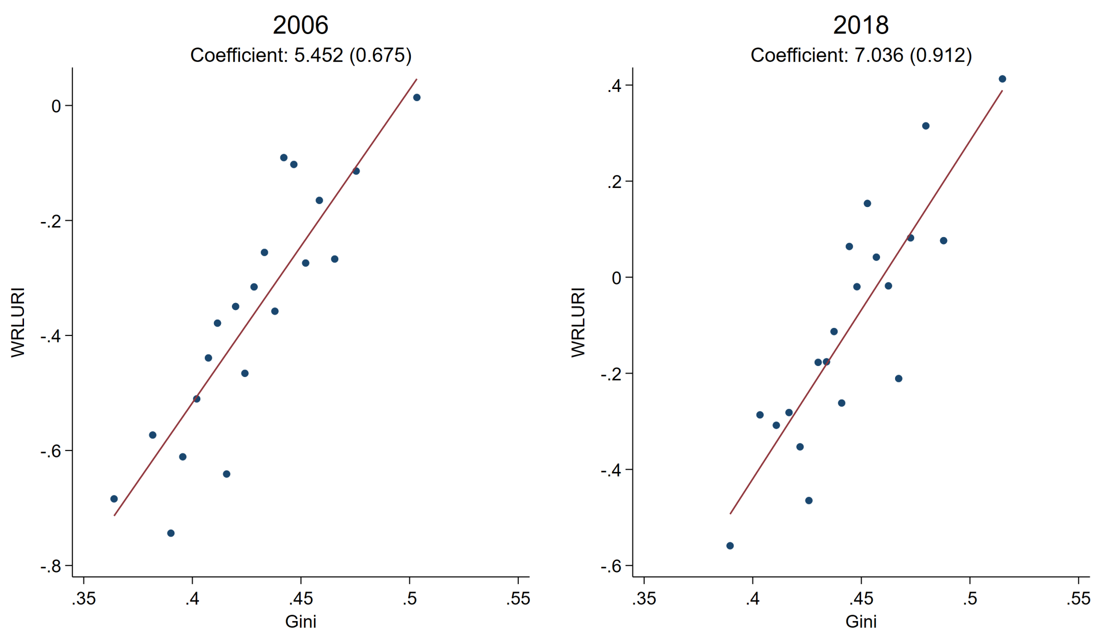


Figure 3. Correlation Between the Gini Coefficient and Other Variables (cont.)

(c) Wharton Residential Land Use Regulation Index



(d) Build Permits

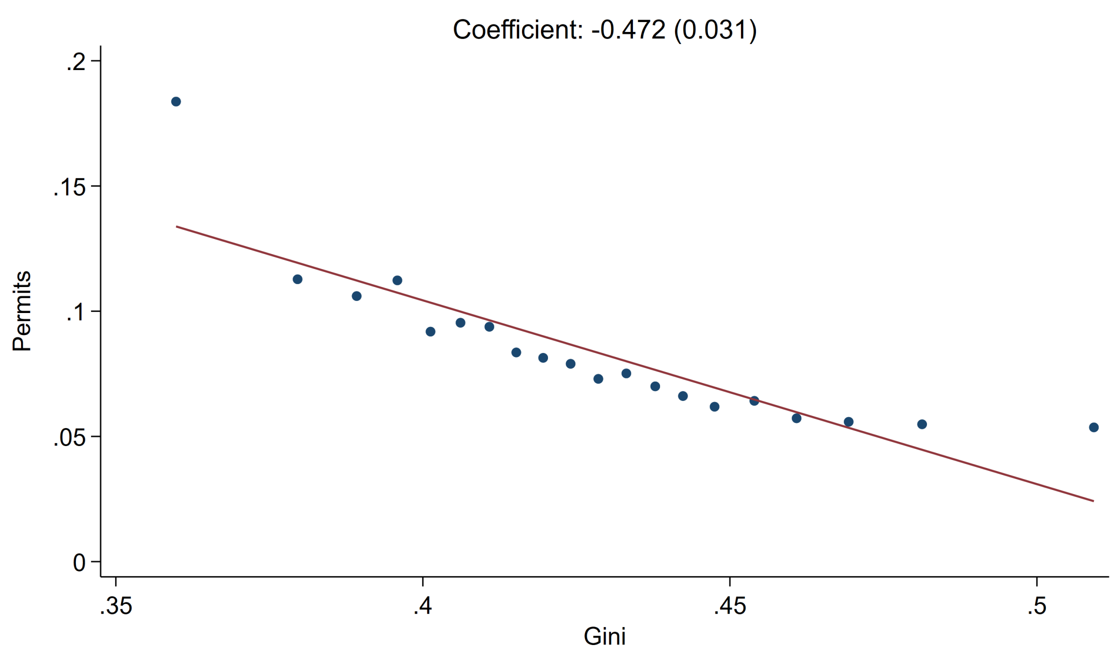
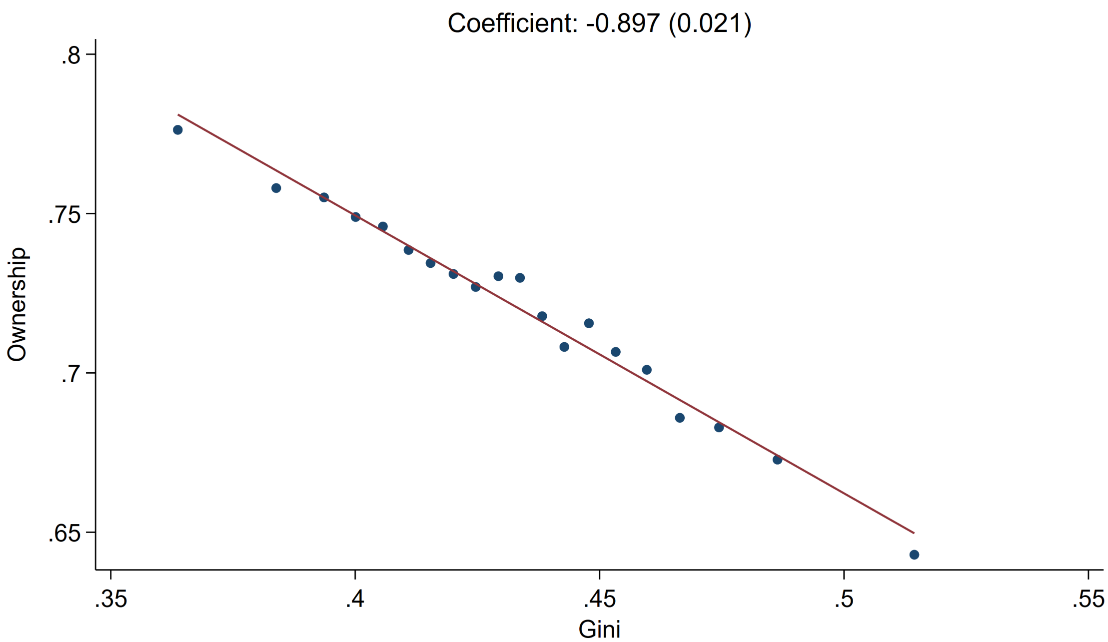


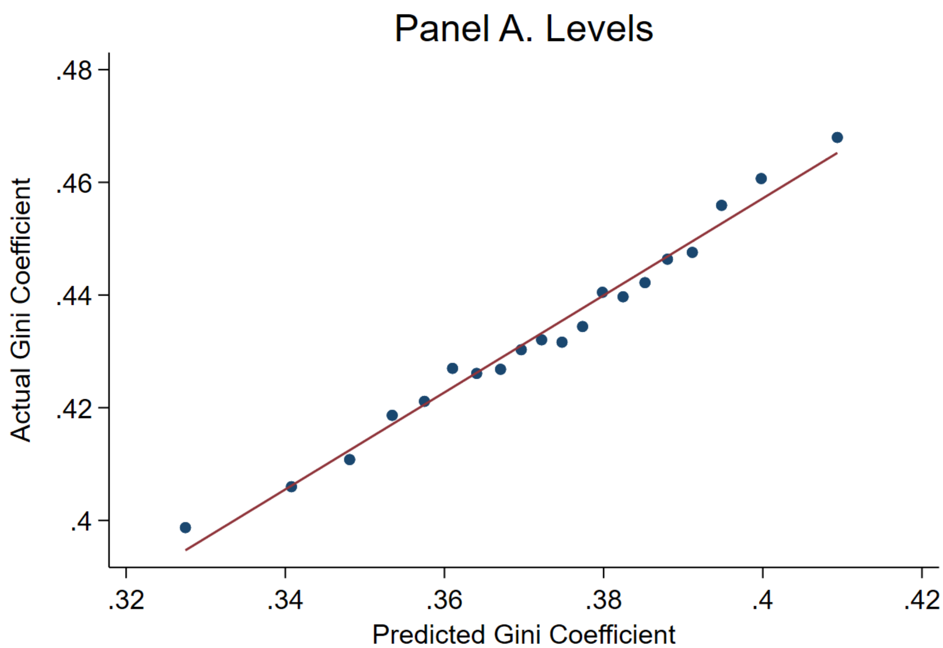
Figure 3. Correlation Between the Gini Coefficient and Other Variables (cont.)

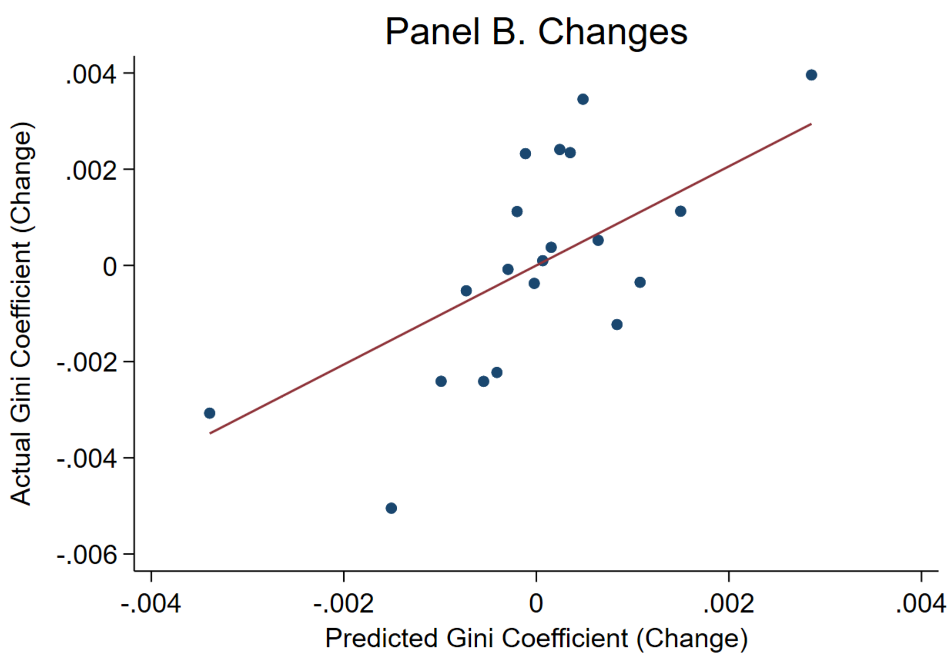
(e) Homeownership



Notes: This figure uses binned scatter plots to show the correlation between the Gini coefficient and various housing market outcomes using the entire pooled sample. The outcomes of interest include the logarithm of median housing value (Panel A), the logarithm of the number of owner-occupied housing units (Panel B), the Wharton Residential Land Use Regulation Index (Panel C), the number of building permits issued over the next 10 years scaled by current housing stocks (Panel D), and the homeownership rate (Panel E). All monetary variables are adjusted to 1990 dollars. The solid line in each panel represents the linear fit of the relationship after controlling for log real mean household income and log population if the outcome is a quantity measure, with the corresponding coefficient and standard error (in parentheses) reported.

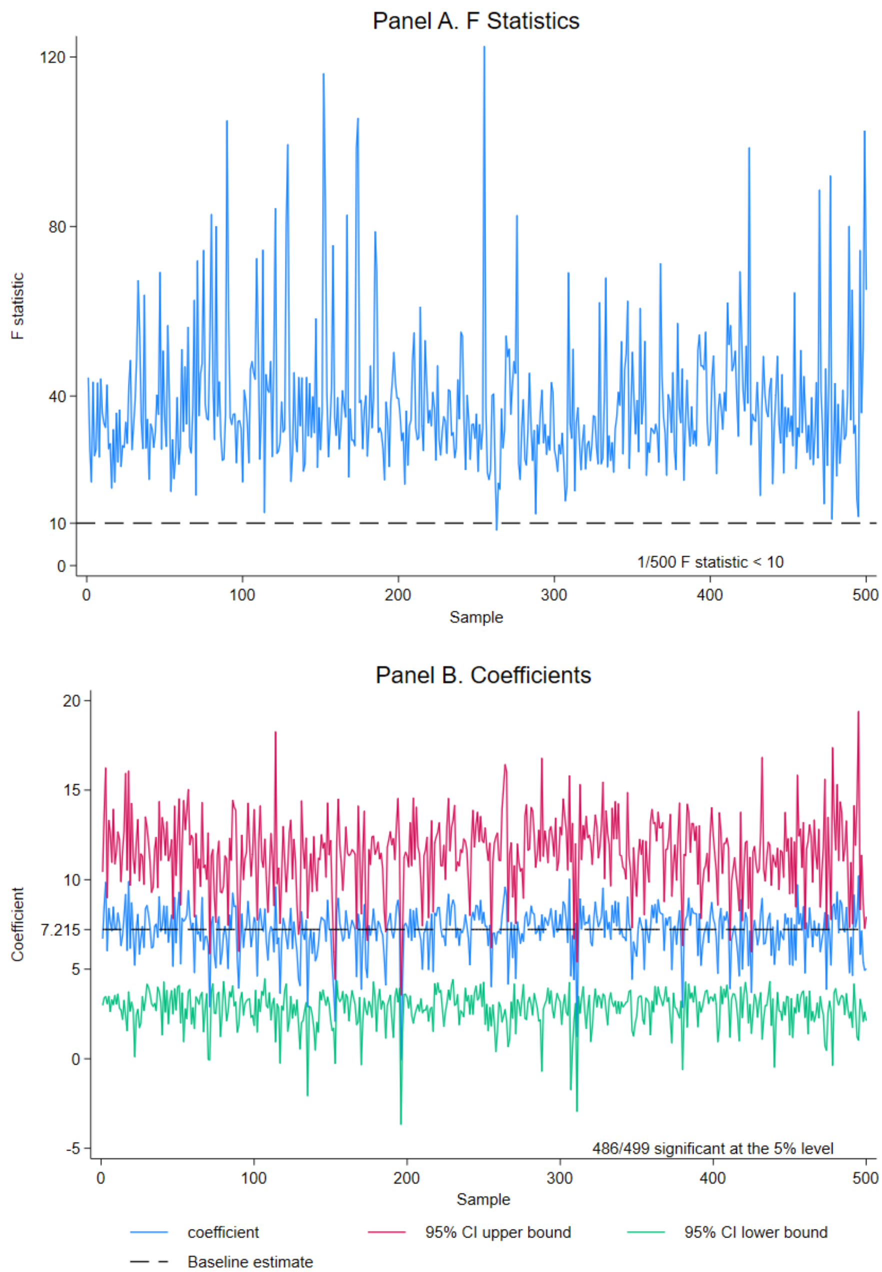
Figure 4. Predicted and Actual Gini Coefficients





Notes: Panel A of this figure is a binned scatter plot of the predicted against actual Gini coefficients. Panel B of this figure is a binned scatter plot of the residual changes in the predicted against actual Gini coefficients over a decade, after controlling for changes in mean income, total population, unemployment rate, the share of minorities, the share of population with post-secondary education, and state-year fixed effects.

Figure 5. Use Random 80 Occupation-Income Groups to Construct IVs



Notes: We randomly select 80 occupation-income groups to construct our IV and re-estimate our model for 500 times. Panel A plots the first-stage F-statistics across 500 iterations. One out of 500 iterations is smaller than ten. Panel B plots the estimated coefficients and their 95% confidence intervals for 499 iterations with firs-stage F-statistic above ten. 486 out of 499 iterations are statistically significant at the 5% level. The black dashed line represents the baseline estimate from column 2 of Table 3.

Table 1. Summary Statistics

Sample	1990 (1)	2000 (2)	2010 (3)	2017 (4)	Total (5)
Gini coefficient	0.415 (0.035)	0.430 (0.037)	0.434 (0.034)	0.445 (0.034)	0.433 (0.036)
Top 20% income share	0.454 (0.029)	0.469 (0.031)	0.472 (0.030)	0.482 (0.031)	0.471 (0.032)
$\overline{\text{5th Quintile}}$ / $\overline{\text{3rd Quintile}}$	2.801 (0.351)	2.993 (0.404)	3.053 (0.406)	3.174 (0.434)	3.034 (0.424)
$\overline{\text{3rd Quintile}}$ / $\overline{\text{1st Quintile}}$	4.073 (0.633)	4.057 (0.664)	4.087 (0.587)	4.204 (0.690)	4.111 (0.648)
log(Building Permits)	7.847 (1.447)	7.096 (1.818)	6.125 (1.989)		6.861 (1.940)
log(Origination)	5.335 (2.235)	6.197 (1.539)	5.432 (1.569)	5.967 (1.574)	5.772 (1.718)
Non-owner-occupied (%)	0.164 (0.215)	0.089 (0.069)	0.139 (0.107)	0.124 (0.093)	0.126 (0.122)
log(Population)	11.155 (1.103)	10.682 (1.204)	10.554 (1.294)	10.561 (1.320)	10.682 (1.268)
log(Median value)	11.053 (0.407)	11.331 (0.369)	11.732 (0.443)	11.880 (0.440)	11.567 (0.517)
log(Housing units)	10.286 (1.080)	9.856 (1.164)	9.785 (1.235)	9.813 (1.245)	9.889 (1.209)
Homeownership (%)	0.698 (0.082)	0.734 (0.078)	0.726 (0.077)	0.715 (0.081)	0.720 (0.080)
log(Mean income)	10.411 (0.196)	10.739 (0.197)	10.974 (0.211)	11.149 (0.219)	10.878 (0.326)
log(Median income)	10.209 (0.218)	10.503 (0.217)	10.717 (0.234)	10.874 (0.241)	10.629 (0.322)
Minorities (%)	0.127 (0.142)	0.145 (0.163)	0.168 (0.174)	0.179 (0.177)	0.159 (0.168)
Unemployed (%)	0.062 (0.021)	0.057 (0.024)	0.088 (0.034)	0.052 (0.022)	0.065 (0.030)
Post-secondary educ. (%)	0.411 (0.106)	0.443 (0.112)	0.497 (0.106)	0.534 (0.105)	0.481 (0.116)
Observations	1,439	2,392	2,725	2,731	9,287

Notes: This table reports summary statistics for key variables at the county level. Columns 1 to 4 present the mean and standard deviation (in parentheses) for each variable in 1990, 2000, 2010, and 2017, respectively. Column 5 reports the overall mean and standard deviation for the pooled sample. All monetary variables are expressed in nominal values. Building permits data are not available for 2017, because the variable is the number of permits issued in the following decade. See Section I for detailed information on data sources and variable construction.

Table 2. House Prices and Income Inequality (Reverse Causality)

Dependent variable	Poverty rate		Top 20% share		Gini	
	OLS	IV	OLS	IV	OLS	IV
	(1)	(2)	(3)	(4)	(5)	(6)
log(HPI)	-0.019**	-0.062**	-0.017***	-0.034***	-0.019***	-0.027**
	(0.008)	(0.026)	(0.003)	(0.009)	(0.003)	(0.011)
Observations	9,198	9,198	9,198	9,198	9,198	9,198
F statistic		52.785		52.785		52.785
County controls	Yes	Yes	Yes	Yes	Yes	Yes
County fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table reports the estimation results of the effect of the log of house price index on income inequality. Dependent variables are the poverty rate (columns 1 and 2), the income share of top 20 percent (columns 3 and 4), and the Gini coefficient (columns 5 and 6). We use the interaction term between the national house price index and local land unavailability as the IV for local house price index. The F statistic reported is the Kleibergen-Paap rk F statistic. County controls include the log of mean income, the log of total population, unemployment rate, the proportion of minority populations, and the percentage of population with post-secondary education. All regressions include county fixed effects and year fixed effects. All regressions are weighted by the number of households for each observation. Standard errors in parentheses are clustered at the county level. Significance levels 10%, 5%, 1% are denoted by *, **, and ***, respectively.

Table 3. Income Inequality and House Prices

Dependent variable	Gini	log(HPI)			
	First-stage (1)	IV-all (2)	IV-inelastic (3)	IV-elastic (4)	OLS-all (5)
Predicted Gini	0.926***				
	(0.153)				
Gini		7.215***	11.136***	3.688	-0.593***
		(2.135)	(4.025)	(2.784)	(0.211)
Observations	9,277	9,194	4,612	4,568	9,194
F statistic		35.569	20.382	9.921	
County controls	Yes	Yes	Yes	Yes	Yes
County fixed effects	Yes	Yes	Yes	Yes	Yes
State-year fixed effects	Yes	Yes	Yes	Yes	Yes

Notes: This table reports the estimation results of the effect of the Gini coefficient on the log of house price index. Column 1 reports the first-stage result. Columns 2 and 5 report the 2SLS and OLS results for the full sample, respectively. Columns 3 and 4 report the 2SLS results for counties with inelastic and elastic land supply, respectively. Inelastic (elastic) counties have a land unavailability index greater than (less than) 21, following Lutz and Sand (2023). The F statistic reported is the Kleibergen-Paap rk F statistic. County controls include the log of mean income, the log of total population, unemployment rate, the proportion of minority populations, and the percentage of population with post-secondary education. All regressions include county fixed effects and state-year fixed effects. All regressions are weighted by the number of households for each observation. Standard errors in parentheses are clustered at the county level. Significance levels 10%, 5%, 1% are denoted by *, **, and ***, respectively.

Table 4. Drop 10 Potentially Influential Occupation-Income Groups (IV Estimates)

Dependent variable	log(HPI)			
	Drop 10 highest (1)	Drop 10 lowest (2)	Drop 10 largest (3)	Drop 5 highest & 5 lowest (4)
Gini	5.233***	7.739***	5.140***	7.234***
	(1.685)	(2.741)	(1.707)	(2.636)
Observations	9,194	9,194	9,194	9,194
F statistic	42.927	25.714	46.889	24.390
County controls	Yes	Yes	Yes	Yes
County fixed effects	Yes	Yes	Yes	Yes
State-year fixed effects	Yes	Yes	Yes	Yes

Notes: This table reports the IV estimation results of the effect of the Gini coefficient on the log of house price index using alternative instruments. We drop different sets of 10 potentially influential occupation-income groups to construct the predicted Gini coefficients. In column 1, we drop 10 groups with the highest income. In column 2, we drop 10 groups with the lowest income. In column 3, we drop 10 groups with the largest employment shares. In column 4, we drop 5 groups with the highest income and 5 groups with the lowest income. The F statistic reported is the Kleibergen-Paap rk F statistic. County controls include the log of mean income, the log of total population, unemployment rate, the proportion of minority populations, and the percentage of population with post-secondary education. All regressions include county fixed effects and state-year fixed effects. All regressions are weighted by the number of households for each observation. Standard errors in parentheses are clustered at the county level. Significance levels 10%, 5%, 1% are denoted by *, **, and ***, respectively.

Table 5. Income Inequality and Housing Stocks

Dependent variable	log(Housing units)		log(Owner occupied units)		Housing units per capita	
	OLS (1)	IV (2)	OLS (3)	IV (4)	OLS (5)	IV (6)
Gini	0.244***	-3.910***	0.032	-3.612***	0.092***	-1.548***
	(0.044)	(0.893)	(0.074)	(0.835)	(0.020)	(0.357)
Observations	9,277	9,277	9,277	9,277	9,277	9,277
F statistic		36.568		36.568		36.568
County controls	Yes	Yes	Yes	Yes	Yes	Yes
County fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
State-year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table reports the estimation results of the effect of the Gini coefficient on housing stocks. Dependent variables are the log of housing units (columns 1 and 2), the log of owner-occupied housing units (columns 3 and 4), and the per capita housing units (columns 5 and 6). The F statistic reported is the Kleibergen-Paap rk F statistic. County controls include the log of mean income, the log of total population, unemployment rate, the proportion of minority populations, and the percentage of population with post-secondary education. All regressions include county fixed effects and state-year fixed effects. All regressions are weighted by the number of households for each observation. Standard errors in parentheses are clustered at the county level. Significance levels 10%, 5%, 1% are denoted by *, **, and ***, respectively.

Table 6. Income Inequality and Housing Regulations

Dependent variable	LPPI		SRI		WRLURI	
	OLS	IV	OLS	IV	OLS	IV
	(1)	(2)	(3)	(4)	(5)	(6)
	Panel A. 2006					
Gini	0.458	15.317**	-2.882	6.599**	-0.111	13.497**
	(2.172)	(5.752)	(2.232)	(3.070)	(1.633)	(5.523)
Observations	825	825	825	825	825	825
F statistic		30.641		30.641		30.641
	Panel B. 2018					
Gini	-0.030	23.296***	-0.865	5.041**	-1.064	10.045**
	(2.263)	(7.208)	(1.657)	(2.142)	(2.390)	(4.012)
Observations	841	841	841	841	841	841
F statistic		20.710		20.710		20.710
County controls	Yes	Yes	Yes	Yes	Yes	Yes
State fixed effects	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table reports the estimation results of the effect of the Gini coefficient on housing regulations. Panels A and B use the 2006 and 2018 Wharton Residential Land Use Regulation Index (WRLURI), respectively. Dependent variables are the Local Political Pressure Index (LPPI) in columns 1 and 2, the Supply Restrictions Index (SRI) in columns 3 and 4, and the WRLURI in columns 5 and 6. The F statistic reported is the Kleibergen-Paap rk F statistic. County controls include the log of mean income, the log of total population, unemployment rate, the proportion of minority populations, the proportion of people over 65, poverty rate, and land unavailability index. All regressions include state fixed effects. All regressions are weighted by the number of households for each observation. Standard errors in parentheses are clustered at the state level. Significance levels 10%, 5%, 1% are denoted by *, **, and ***, respectively.

Table 7. Income Inequality and Building Permits

Dependent variable	log(Permits)		Permits / Population		Multifamily permits / Population	
	OLS	IV	OLS	IV	OLS	IV
	(1)	(2)	(3)	(4)	(5)	(6)
Gini	1.689	-16.150**	-0.005	-0.492**	-0.002*	-0.022***
	(1.096)	(6.827)	(0.034)	(0.244)	(0.001)	(0.008)
Observations	6,125	6,125	6,147	6,147	6,147	6,147
F statistic		37.242		37.110		37.110
County controls	Yes	Yes	Yes	Yes	Yes	Yes
County fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
State-year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table reports the estimation results of the effect of the Gini coefficient on housing supply. Dependent variables are the log of building permits issued over the following decade (columns 1 and 2), the per capita building permits issued over the following decade (columns 3 and 4), and the per capita building permits issued for projects with more than 5 units over the following decade (columns 5 and 6). The F statistic reported is the Kleibergen-Paap rk F statistic. County controls include the log of mean income, the log of total population, unemployment rate, the proportion of minority populations, and the percentage of population with post-secondary education. All regressions include county fixed effects and state-year fixed effects. All regressions are weighted by the number of households for each observation. Standard errors in parentheses are clustered at the county level. Significance levels 10%, 5%, 1% are denoted by *, **, and ***, respectively.

Table 8. Income Inequality and Homeownership Rate

Dependent variable	Homeownership rate			
	OLS-all (1)	IV-all (2)	IV-inelastic (3)	IV-elastic (4)
Gini	-0.143*** (0.031)	-0.521** (0.250)	-1.707*** (0.408)	0.344 (0.478)
Observations	9,277	9,277	4,654	4,609
F statistic		36.568	20.954	10.250
County controls	Yes	Yes	Yes	Yes
County fixed effects	Yes	Yes	Yes	Yes
State-year fixed effects	Yes	Yes	Yes	Yes

Notes: This table reports the estimation results of the effect of the Gini coefficient on homeownership rates. Column 1 reports the first-stage result. Columns 2 reports the 2SLS result for the full sample. Columns 3 and 4 report the 2SLS results for counties with inelastic and elastic land supply, respectively. Inelastic (elastic) counties have a land unavailability index greater than (less than) 21, following Lutz and Sand (2023). The F statistic reported is the Kleibergen-Paap rk F statistic. County controls include the log of mean income, the log of total population, unemployment rate, the proportion of minority populations, and the percentage of population with post-secondary education. All regressions include county fixed effects and state-year fixed effects. All regressions are weighted by the number of households for each observation. Standard errors in parentheses are clustered at the county level. Significance levels 10%, 5%, 1% are denoted by *, **, and ***, respectively.

Income Inequality, House Prices, and Housing Regulations
Online Appendix

A Construction of Instrumental Variable for Income Inequality

We construct a Bartik-style instrumental variable for the Gini coefficient. Although this IV is not in the canonical form of traditional Bartik IVs, it has an initial share component and a national shift component. Our approach involves the following steps:

1. Freeze the shares of occupation-income groups at their 1980 levels for each county.
2. Calculate the national median income for each occupation-income group in later years, excluding data from the state being considered (leave-one-out strategy).
3. Predict the income distribution for each county by assuming that the income level of each occupation-income group grows according to the nationwide change, while the shares remain fixed at their 1980 levels.

1 A Concrete Example

We use Cook county in Illinois as an example to walk you through the construction procedure of our instrument.

Step 1a: We start with the county-level employment shares across 15 occupations in 1980. For the sake of convenience in presentation, here we use their population counts.

Occupation	Population
Executive and managerial occupations	257,626
Professional specialty occupations	289,086
Administrative occupations	509,018
...	...
Military	2,704
Unemployed	189,937

For example, in 1980, Cook county had 257,626 individuals employed in executive and managerial occupations.

Step 1b: We exploit the national income distribution for each occupation in 1980 to further divide the state-level occupation employment into 6 income percentile bins.

State	Occupation	Group	Percentile	Share
IL	Executive and managerial	1	0-10	0.07
IL	Executive and managerial	2	10-30	0.16
IL	Executive and managerial	3	30-70	0.40
IL	Executive and managerial	4	70-90	0.24
IL	Executive and managerial	5	90-98	0.06
IL	Executive and managerial	6	98-100	0.07

As indicated in the table's final row, 7% of people in Illinois with executive and managerial occupations earned incomes above the 98th percentile of the national income distribution for this occupation. The percentile thresholds are the same across all states, but the resulting distributions of these 6 income percentile bins can vary. With numerous large firm headquarters located in Chicago, the share of 98-100 percentile bin exceeds 2%.

Step 1c: We combine the above two steps to generate the initial shares of 90 occupation-income groups in 1980.

State	Occupation	Group	Percentile	Population
IL	Executive and managerial	1	0-10	257,626×0.07
IL	Executive and managerial	2	10-30	257,626×0.16
IL	Executive and managerial	3	30-70	257,626×0.40
IL	Executive and managerial	4	70-90	257,626×0.24
IL	Executive and managerial	5	90-98	257,626×0.06
IL	Executive and managerial	6	98-100	257,626×0.07

In our example, we are assuming that about 18,034 households in an occupation-income group identified by executive and managerial positions and earning above the 98th percentile.

Step 2: We use a leave-one-out approach to determine national income levels for each occupation-income groups in subsequent years. Here we use the year 1990 as an example.

State	Occupation	p5	p20	p50	p80	p94	p99
IL	Executive occupations	16,000	31,001	53,763	86,852	139,978	254,000
IL	Professional occupations	15,000	30,010	51,000	80,362	126,900	229,716
IL	Administrative occupations	11,626	22,728	40,000	62,494	91,190	154,700
...	...	...	...	...	...	...	...
IL	Military	12,100	18,918	30,220	48,733	72,000	109,000
IL	Unemployed	2,500	10,000	25,200	48,200	76,400	129,603

To calculate the 99th percentile income for executive and managerial occupations, we use micro-level census data excluding Illinois observations from 1990. This process involves identifying the midpoint income for each percentile bin, such as using the 99th percentile income as a benchmark for the 98-100th percentile bin.

Step 3: We project an income distribution for each county for the year 1990.

Group	Occupation	Percentile	Income	Household
1	Executive occupations	0-10	16,000	257,626×0.07
2	Executive occupations	10-30	31,001	257,626×0.16
3	Executive occupations	30-70	53,763	257,626×0.40
4	Executive occupations	70-90	86,852	257,626×0.24
5	Executive occupations	90-98	139,978	257,626×0.06
6	Executive occupations	98-100	254,000	257,626×0.07

Our method forecasts that 18,034 households in Cook county are predicted to have an income of $254,000 in 1990. Then we can use Equation (3) to calculate the predicted Gini coefficient.

2 Comparison with Boustan et al. (2013)

Our approach to construct the predicted Gini coefficient draws inspiration from Boustan et al. (2013), yet with several key modifications to their original methodology.

The main difference arises in the formulation of initial shares. They use the initial (1970) household counts across 15 income bins based on the 1970 census, ranging from under $1,000 to $50,000 and over. They then convert the end points of each income bin from absolute income levels into percentiles of the income distribution. For example, the first income bin includes households earning up to $1,000 or up to the 3.7th percentile of the income distribution in 1970. Therefore, their initial shares are essentially the 15 shares of households within each income percentile bins.

In contrast, our methodology segments initial shares into 90 occupation-income groups, offering a more detailed snapshot of the initial income distribution. This granularity provides two distinct advantages. Firstly, it enhances the predictive accuracy of future income distributions. We find a weak first-stage correlation between the predicted and actual Gini coefficients when following their approach, as they only have 15 income bins to simulate future income distributions, resulting in predicted Gini coefficients that do not accurately match real Gini coefficients. Second, our initial shares are more likely to satisfy the exclusion restriction, following the interpretation of Bartik-style instruments in Goldsmith-Pinkham, Sorkin and Swift (2020). Given that income inequality is highly correlated over time within the same county, the initial shares of income bins may have a direct effect on future housing prices. However, we argue that some of our initial shares are more likely to be exogenous. For instance, it seems implausible that the share of military workers in the 10th to 30th income percentiles in 1980 could directly affect future house prices, as military employment is largely determined by national defense policies and is unlikely to be correlated with local housing market conditions. Moreover, we use two tests to show that our 90 initial shares contribute similarly to the overall identification.

In addition to the difference in initial shares, we use a slightly different method to construct the national shifts. Specifically, we apply a leave-one-out strategy for each state when calculating the national median income of each occupation-income group. This approach helps to address the potential finite sample bias that may arise from using a state's own-observation information.

Table A1. Occupations and OCC1990 Codes

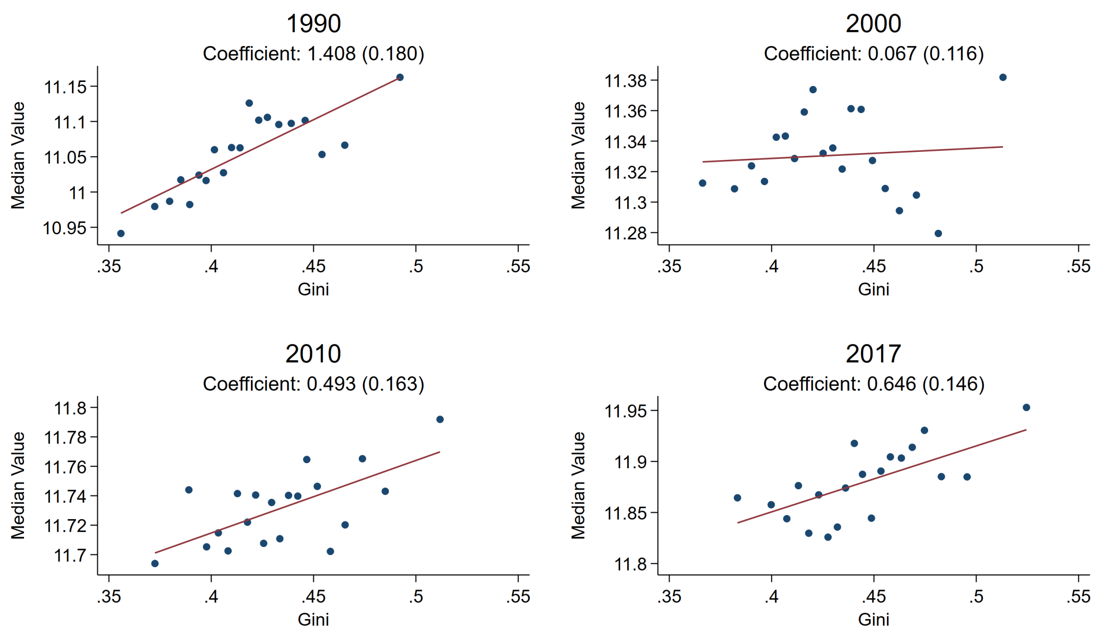
Occupation	OCC1990
Executive, Administrative, and Managerial Occupations	3-22
Management-Related Occupations and Professional Specialty Occupations	23-200
Technicians and Related Support Occupations	203-235
Sales Occupations	243-283
Administrative Support Occupations	303-389
Private Household Occupations	405-407
Protective Service Occupations	415-427
Other Service Occupations	434-469
Farming, Forestry, and Fishing Occupations	473-498
Precision Production, Craft, and Repair Occupations	503-699
Machine Operators, Assemblers, and Inspectors	703-799
Transportation Occupations	803-859
Helpers, Construction and Extractive Occupations	865-889
Military Occupations	905
Unemployed	N.A.

Notes: This table lists the 15 occupations used to construct the initial share component of our instrumental variable, along their corresponding occupation codes (variable *OCC1990*) from the IPUMS U.S. Census 1980 data.

B Additional Figures and Tables

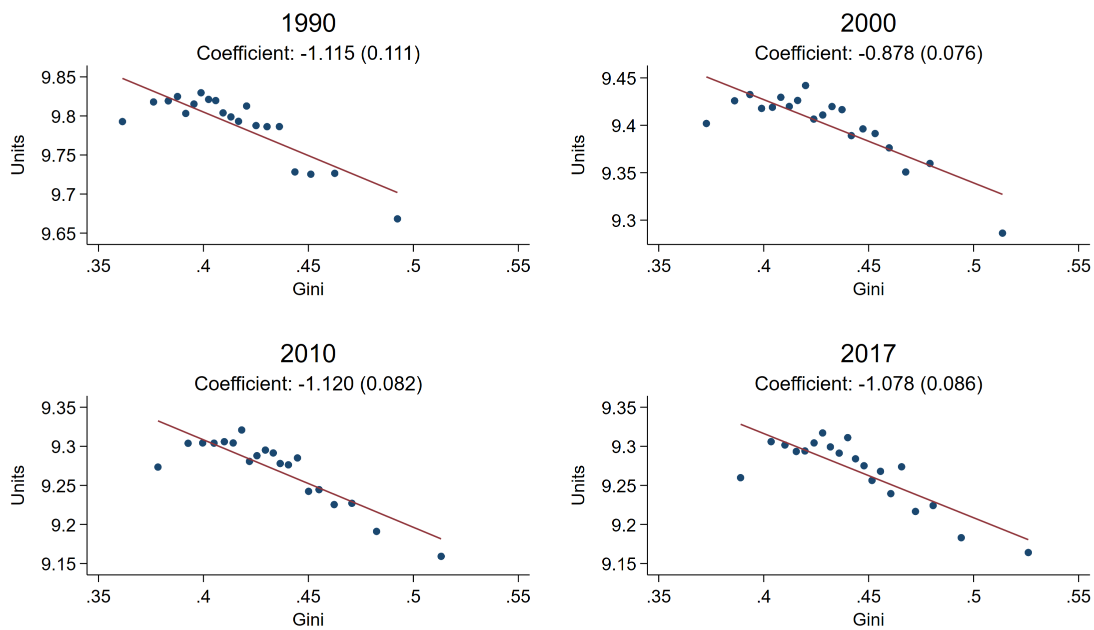
1 Additional Figures

Figure B1. Gini Coefficient and Median Housing Value



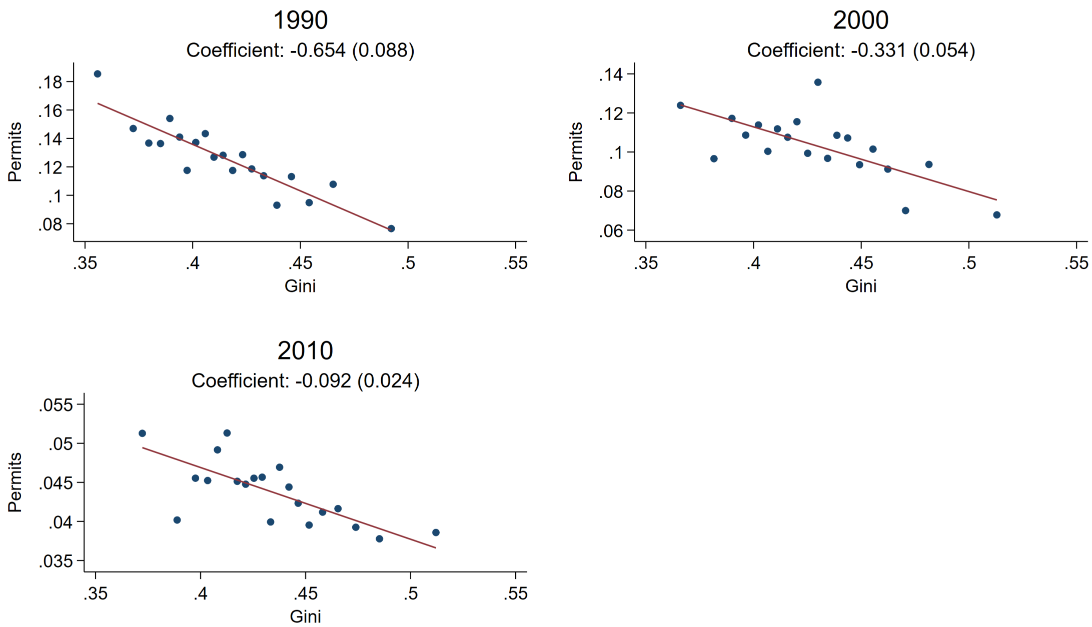
Notes: This figure plots the cross-sectional relationship between the Gini coefficient and the logarithm of median housing value for 1990, 2000, 2010, and 2017. The figure uses binned scatter plots to visualize the relationship, controlling for the logarithm of mean household income. The solid line represents the linear fit of the relationship between the Gini coefficient and log median housing value, with the corresponding coefficient and standard error (in parentheses) reported.

Figure B2. Gini Coefficient and Owner-Occupied Units



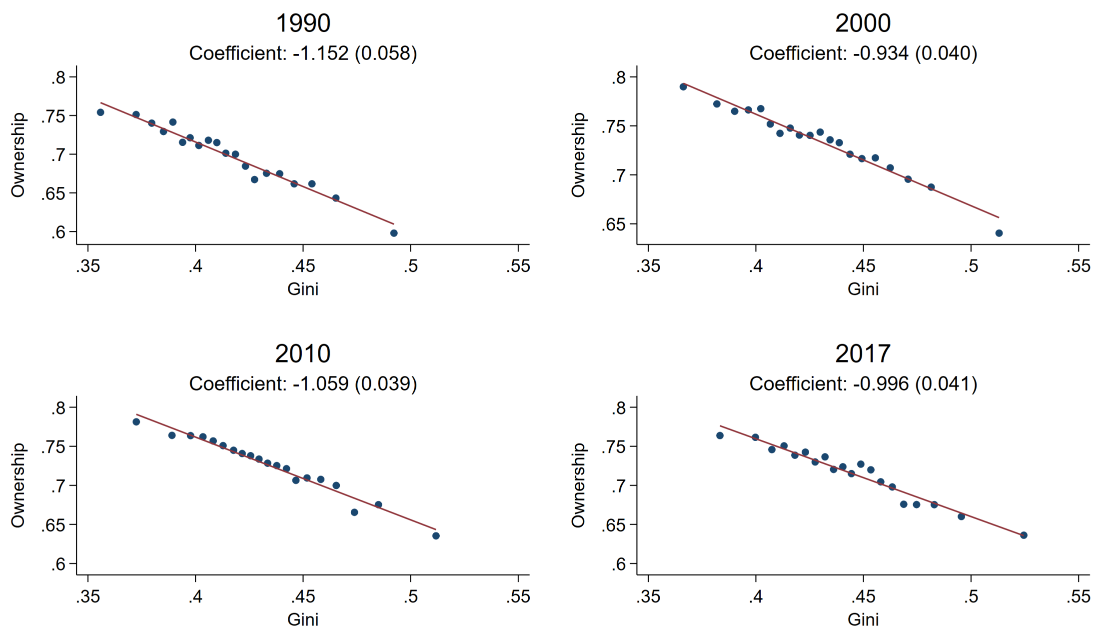
Notes: This figure plots the cross-sectional relationship between the Gini coefficient and the logarithm of the number of owner-occupied housing units for 1990, 2000, 2010, and 2017. The figure uses binned scatter plots to visualize the relationship, controlling for the logarithm of mean household income and population. The solid line represents the linear fit of the relationship between the Gini coefficient and log owner-occupied housing units, with the corresponding coefficient and standard error (in parentheses) reported.

Figure B3. Gini Coefficient and Building Permits



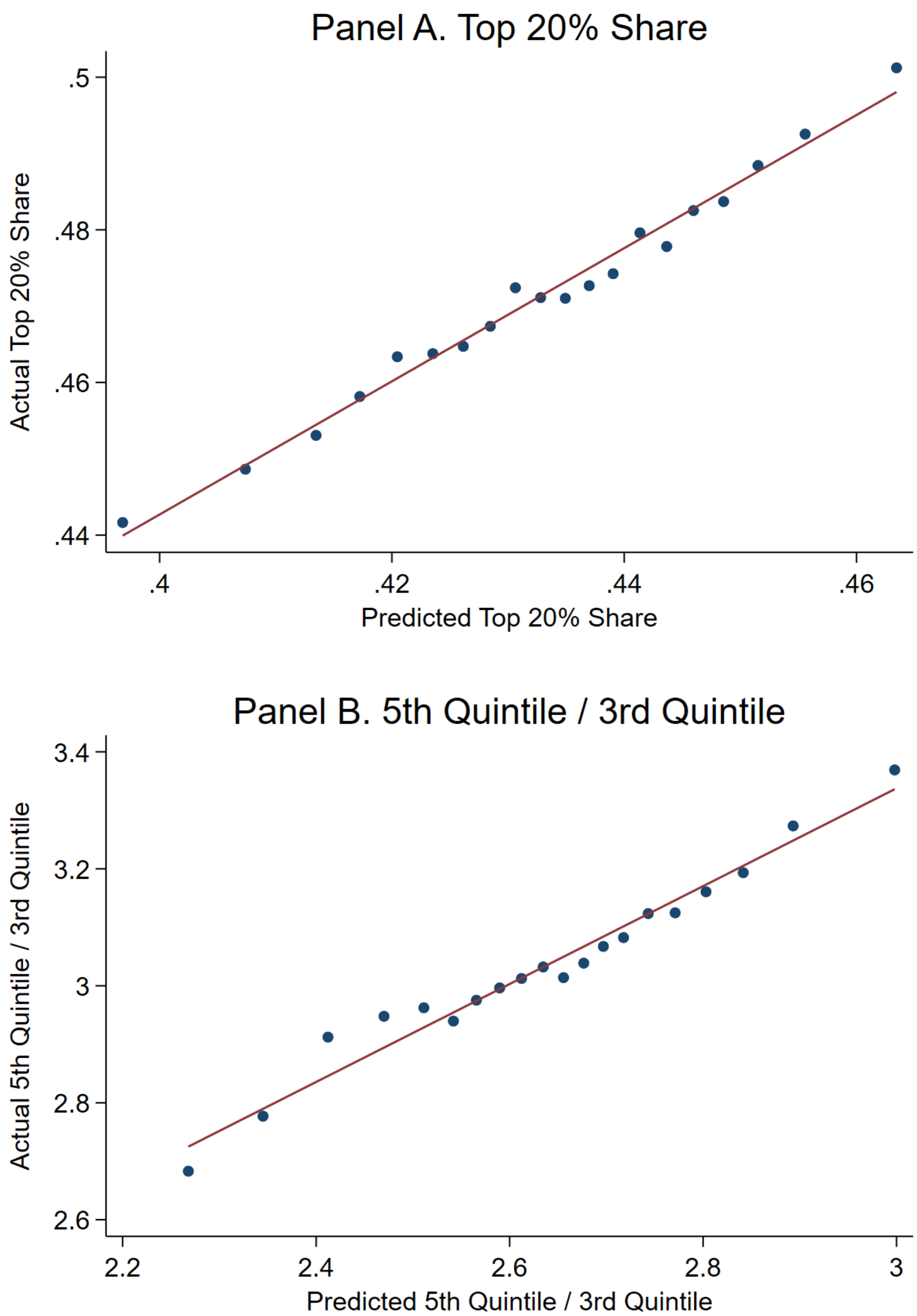
Notes: This figure plots the cross-sectional relationship between the Gini coefficient and the number of building permits issued over the next ten years scaled by current housing stocks for 1990, 2000, and 2010. The figure uses binned scatter plots to visualize the relationship, controlling for the logarithm of mean household income. The solid line represents the linear fit of the relationship between the Gini coefficient and the growth of building permits, with the corresponding coefficient and standard error (in parentheses) reported.

Figure B4. Gini Coefficient and Homeownership Rate



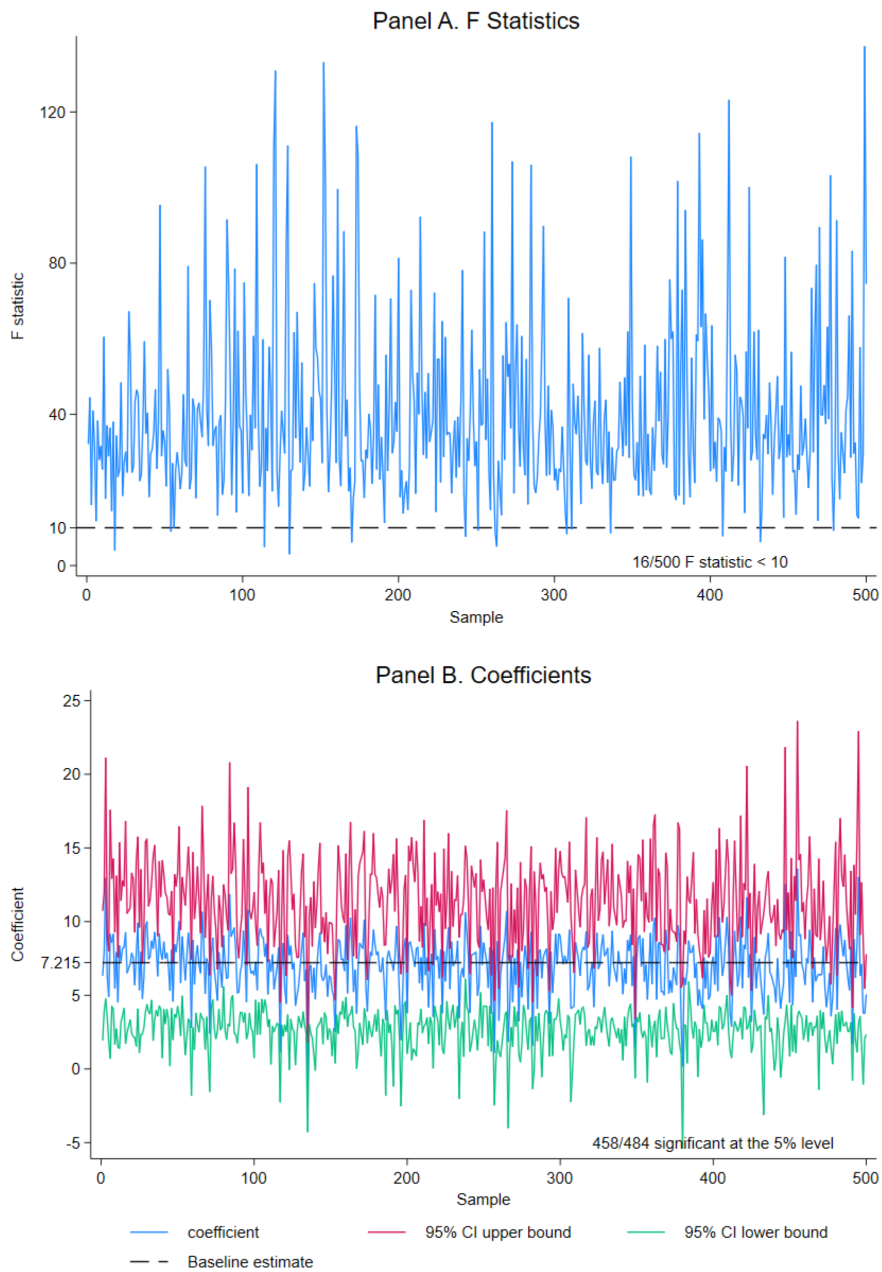
Notes: This figure plots the cross-sectional relationship between the Gini coefficient and homeownership rate for 1990, 2000, 2010, and 2017. The figure uses binned scatter plots to visualize the relationship, controlling for the logarithm of mean household income. The solid line represents the linear fit of the relationship between the Gini coefficient and homeownership rate, with the corresponding coefficient and standard error (in parentheses) reported.

Figure B5. Predicted and Actual Income Inequality Measures



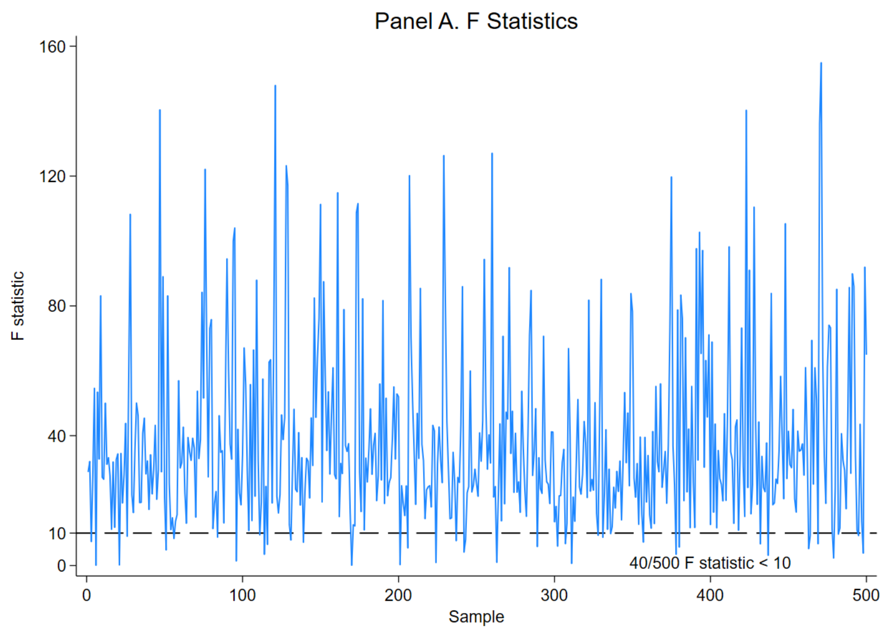
Notes: Panel A of this figure is a binned scatter plot of the predicted against actual income share of the top 20 percent. Panel B of this figure is a binned scatter plot of the predicted against actual ratio between mean income of the highest and third quintiles.

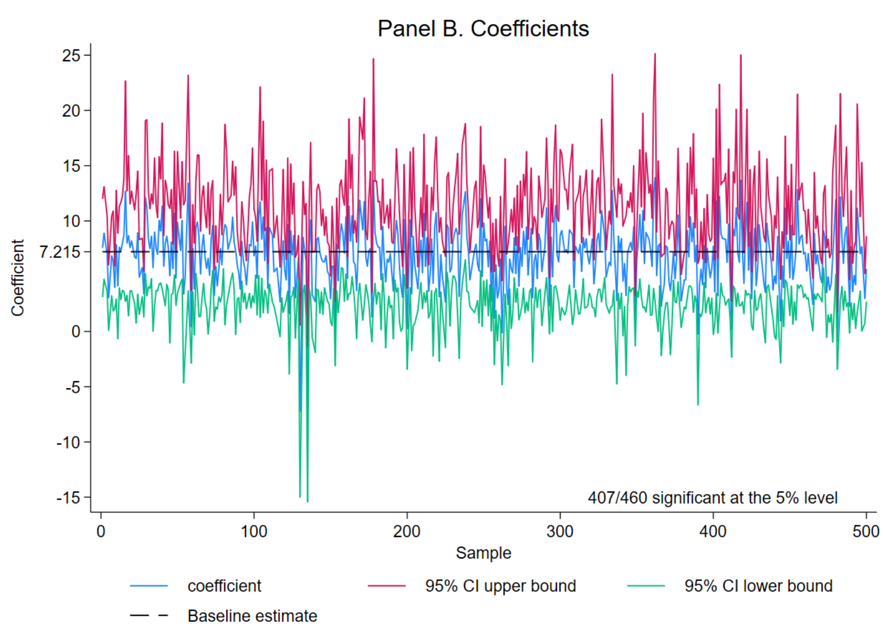
Figure B6. Use Random 70 Occupation-Income Groups to Construct IVs



Notes: We randomly select 70 occupation-income groups to construct our IV and re-estimate our model for 500 times. Panel A plots the first-stage F-statistics across 500 iterations. 16 out of 500 iterations is smaller than ten. Panel B plots the estimated coefficients and their 95% confidence intervals for 484 iterations with firs-stage F-statistic above ten. 458 out of 484 iterations are statistically significant at the 5% level. The black dashed line represents the baseline estimate from column 2 of Table 3.

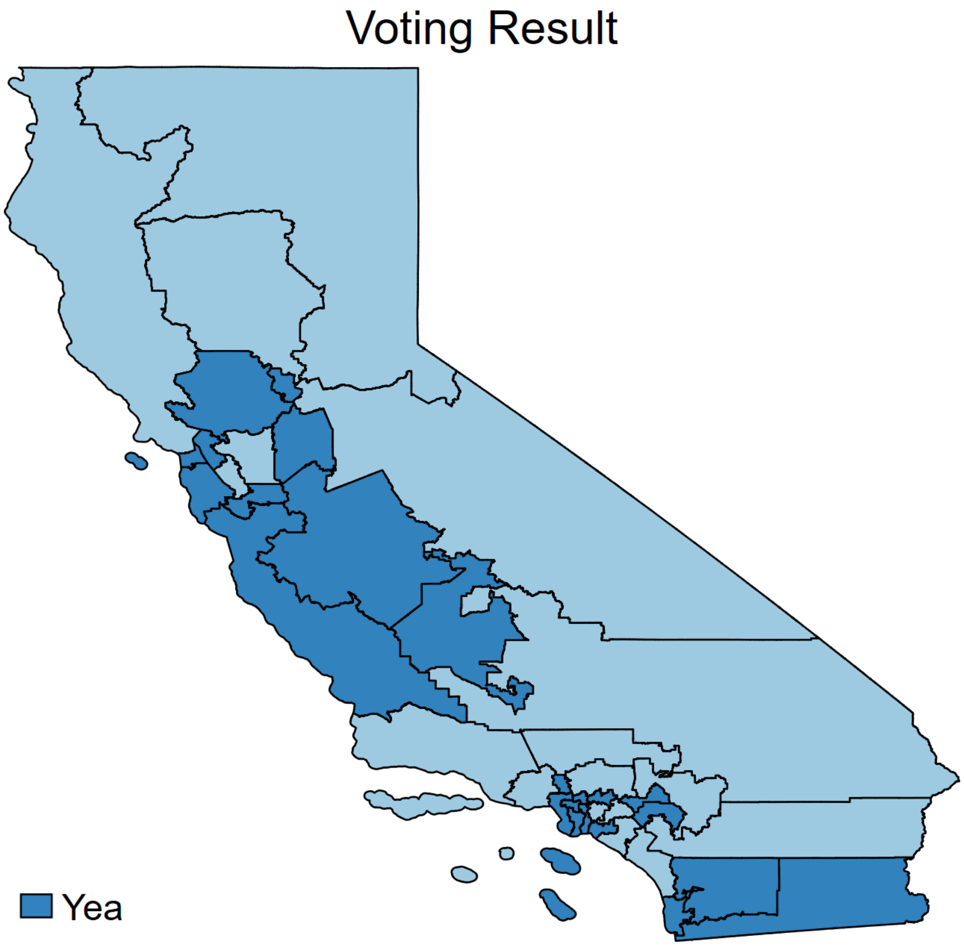
Figure B7. Use Random 60 Occupation-Income Groups to Construct IVs





Notes: We randomly select 60 occupation-income groups to construct our IV and re-estimate our model for 500 times. Panel A plots the first-stage F-statistics across 500 iterations. 40 out of 500 iterations is smaller than ten. Panel B plots the estimated coefficients and their 95% confidence intervals for 460 iterations with firs-stage F-statistic above ten. 407 out of 460 iterations are statistically significant at the 5% level. The black dashed line represents the baseline estimate from column 2 of Table 3.

Figure B8. Voting Results of California Senate Bill 35



Notes: This figure plots the voting results of California Senate Bill 35, with senate districts that voted "Yea" (or in favor) represented in darker blue.

2 Additional Tables

Table B1. House Prices and Its Instrumental Variable (First-Stage)

Dependent variable		log(HPI)	
	(1)	(2)	(3)
log(National HPI) × Land unavailability	0.001***	0.001***	0.004***
	(0.0003)	(0.0002)	(0.0005)
Observations	9,230	9,230	9,198
County controls	No	Yes	Yes
County fixed effects	No	No	Yes
Year fixed effects	No	No	Yes

Notes: This table reports the estimation results of the relationship between the log of local house price index and its instrumental variable, which is the interaction term between the national house price index and local land unavailability. County controls include the log of mean income, the log of total population, unemployment rate, the proportion of minority populations, and the percentage of population with post-secondary education. All regressions are weighted by the number of households for each observation. Standard errors in parentheses are clustered at the county level. Significance levels 10%, 5%, 1% are denoted by *, **, and ***, respectively.

Table B2. Income Inequality and House Prices (Specification of Kösem (2023))

Dependent variable	log(HPI)			
	(1)	(2)	(3)	(4)
Gini	-1.612***	-0.412**		
	(0.414)	(0.204)		
Top 5% share			-1.936***	-0.714***
			(0.387)	(0.175)
Observations	8,927	8,923	8,927	8,923
County controls	Yes	Yes	Yes	Yes
County fixed effects	Yes	Yes	Yes	Yes
Year fixed effects	Yes	No	Yes	No
State-year fixed effects	No	Yes	No	Yes

Notes: This table reports the estimation results of the relationship between inocme inequality and the log of house price index, following the specification from Kösem (2023). We use 2 measures of income inequality: the Gini coefficient and the income share of the top 5 percent. County controls include the log of mean income, the log of total population, homeownership rate, and the log of building permits issued. Columns 1 and 3 include county fixed effects and year fixed effects. Columns 2 and 4 include county fixed effects and state-year fixed effects. All regressions are weighted by the number of households for each observation. Standard errors in parentheses are clustered at the county level. Significance levels 10%, 5%, 1% are denoted by *, **, and ***, respectively.

Table B3. Income Inequality and House Prices (OLS Estimates)

Dependent variable	log(HPI)				
	(1)	(2)	(3)	(4)	(5)
Gini	-0.593***				
	(0.211)				
Top 5% share		-0.828***			
		(0.213)			
Top 20% share			-0.550**		
			(0.261)		
$\overline{\text{5th Quintile}}$ / $\overline{\text{3rd Quintile}}$				-0.008	
				(0.023)	
$\overline{\text{5th Quintile}}$ / $\overline{\text{1st Quintile}}$					-0.001
					(0.002)
Observations	9,194	9,194	9,194	9,194	9,194
County controls	Yes	Yes	Yes	Yes	Yes
County fixed effects	Yes	Yes	Yes	Yes	Yes
State-year fixed effects	Yes	Yes	Yes	Yes	Yes

Notes: This table reports the estimation results of the relationship between income inequality and the log of house price index. We use 5 measures of income inequality: the Gini coefficient, the income share of the top 5 percent, the income share of the top 20 percent, the ratio between the mean income of the highest and third quintiles, and the ratio between the mean income of the highest and lowest quintiles. County controls include the log of mean income, the log of total population, unemployment rate, the proportion of minority populations, and the percentage of population with post-secondary education. All regressions include county fixed effects and state-year fixed effects. All regressions are weighted by the number of households for each observation. Standard errors in parentheses are clustered at the county level. Significance levels 10%, 5%, 1% are denoted by *, **, and ***, respectively.

Table B4. Income Inequality and House Prices (OLS Estimates with Median Income)

Dependent variable	log(HPI)				
	(1)	(2)	(3)	(4)	(5)
Gini	1.056***				
	(0.212)				
Top 5% share		0.789***			
		(0.192)			
Top 20% share			1.472***		
			(0.252)		
$\overline{\text{5th Quintile}}$ / $\overline{\text{3rd Quintile}}$				0.152***	
				(0.022)	
$\overline{\text{5th Quintile}}$ / $\overline{\text{1st Quintile}}$					0.007***
					(0.002)
Observations	9,194	9,194	9,194	9,194	9,194
Median income	Yes	Yes	Yes	Yes	Yes
Other county controls	Yes	Yes	Yes	Yes	Yes
County fixed effects	Yes	Yes	Yes	Yes	Yes
State-year fixed effects	Yes	Yes	Yes	Yes	Yes

Notes: This table reports the estimation results of the relationship between income inequality and the log of house price index, controlling for median household income. We use 5 measures of income inequality: the Gini coefficient, the income share of the top 5 percent, the income share of the top 20 percent, the ratio between the mean income of the highest and third quintiles, and the ratio between the mean income of the highest and lowest quintiles. County controls include the log of median income, the log of total population, unemployment rate, the proportion of minority populations, and the percentage of population with post-secondary education. All regressions include county fixed effects and state-year fixed effects. All regressions are weighted by the number of households for each observation. Standard errors in parentheses are clustered at the county level. Significance levels 10%, 5%, 1% are denoted by *, **, and ***, respectively.

Table B5. Income Inequality and House Prices (Top 20% Share)

Dependent variable	Top 20% share	log(HPI)		
	First-stage (1)	IV-all (2)	IV-inelastic (3)	IV-elastic (4)
Predicted top 20% share	0.682***			
	(0.135)			
Top 20% share		9.998***	12.022**	6.817
		(3.478)	(5.155)	(5.004)
Observations	9,277	9,194	4,612	4,568
F statistic		24.474	14.380	6.588
County controls	Yes	Yes	Yes	Yes
County fixed effects	Yes	Yes	Yes	Yes
State-year fixed effects	Yes	Yes	Yes	Yes

Notes: This table reports the estimation results of the effect of the income share of top 20 percent on the log of house price index. Column 1 reports the first-stage result. Column 2 reports the 2SLS result for the full sample. Columns 3 and 4 report the 2SLS results for counties with inelastic and elastic land supply, respectively. Inelastic (elastic) counties have a land unavailability index greater than (less than) 21, following Lutz and Sand (2023). The F statistic reported is the Kleibergen-Paap rk F statistic. County controls include the log of mean income, the log of total population, unemployment rate, the proportion of minority populations, and the percentage of population with post-secondary education. All regressions include county fixed effects and state-year fixed effects. All regressions are weighted by the number of households for each observation. Standard errors in parentheses are clustered at the county level. Significance levels 10%, 5%, 1% are denoted by *, **, and ***, respectively.

Table B6. Income Inequality and House Prices ($\overline{\text{5th Quintile}}$ / $\overline{\text{3rd Quintile}}$)

Dependent variable	$\frac{\overline{\text{5th Quintile}}}{\overline{\text{3rd Quintile}}}$	log(HPI)		
	First-stage (1)	IV-all (2)	IV-inelastic (3)	IV-elastic (4)
Predicted $\frac{\overline{\text{5th Quintile}}}{\overline{\text{3rd Quintile}}}$	0.894***			
	(0.123)			
$\frac{\overline{\text{5th Quintile}}}{\overline{\text{3rd Quintile}}}$		0.532***	0.787***	0.278*
		(0.118)	(0.214)	(0.144)
Observations	9,277	9,194	4,612	4,568
F statistic		57.685	25.289	36.008
County controls	Yes	Yes	Yes	Yes
County fixed effects	Yes	Yes	Yes	Yes
State-year fixed effects	Yes	Yes	Yes	Yes

Notes: This table reports the estimation results of the effect of the ratio between mean income of the highest and third quintiles on the log of house price index. Column 1 reports the first-stage result. Column 2 reports the 2SLS result for the full sample. Columns 3 and 4 report the 2SLS results for counties with inelastic and elastic land supply, respectively. Inelastic (elastic) counties have a land unavailability index greater than (less than) 21, following Lutz and Sand (2023). The F statistic reported is the Kleibergen-Paap rk F statistic. County controls include the log of mean income, the log of total population, unemployment rate, the proportion of minority populations, and the percentage of population with post-secondary education. All regressions include county fixed effects and state-year fixed effects. All regressions are weighted by the number of households for each observation. Standard errors in parentheses are clustered at the county level. Significance levels 10%, 5%, 1% are denoted by *, **, and ***, respectively.

Table B7. Income Inequality and House Prices (Excluding 2010 Data)

Dependent variable	Gini	log(HPI)			
	First-stage (1)	OLS-all (2)	IV-all (3)	IV-inelastic (4)	IV-elastic (5)
Predicted Gini	1.049***				
	(0.172)				
Gini		-0.403	8.280***	13.477***	3.988
		(0.266)	(2.244)	(4.469)	(2.865)
Observations	6,220	6,183	6,183	3,111	3,061
F statistic			35.807	19.088	12.541
County controls	Yes	Yes	Yes	Yes	Yes
County fixed effects	Yes	Yes	Yes	Yes	Yes
State-year fixed effects	Yes	Yes	Yes	Yes	Yes

Notes: This table reports the estimation results of the effect of the Gini coefficient on the log of house price index, excluding the data from year 2010. Column 1 reports the first-stage result. Columns 2 and 3 reports the OLS and 2SLS result for the full sample, respectively. Columns 3 and 4 report the 2SLS results for counties with inelastic and elastic land supply, respectively. Inelastic (elastic) counties have a land unavailability index greater than (less than) 21, following Lutz and Sand (2023). The F statistic reported is the Kleibergen-Paap rk F statistic. County controls include the log of mean income, the log of total population, unemployment rate, the proportion of minority populations, and the percentage of population with post-secondary education. All regressions include county fixed effects and state-year fixed effects. All regressions are weighted by the number of households for each observation. Standard errors in parentheses are clustered at the county level. Significance levels 10%, 5%, 1% are denoted by *, **, and ***, respectively.

Table B8. Income Inequality and House Prices (Balanced Panel)

Dependent variable	Gini	log(HPI)			
	First-stage (1)	OLS-all (2)	IV-all (3)	IV-inelastic (4)	IV-elastic (5)
Predicted Gini	0.952***				
	(0.162)				
Gini		-0.496*	7.488***	12.477***	3.163
		(0.269)	(2.170)	(4.428)	(2.593)
Observations	5,752	5,731	5,731	2,828	2,883
F statistic			34.212	18.197	10.173
County controls	Yes	Yes	Yes	Yes	Yes
County fixed effects	Yes	Yes	Yes	Yes	Yes
State-year fixed effects	Yes	Yes	Yes	Yes	Yes

Notes: This table reports the estimation results of the effect of the Gini coefficient on the log of house price index in a balanced panel. Column 1 reports the first-stage result. Columns 2 and 3 reports the OLS and 2SLS result for the full sample, respectively. Columns 3 and 4 report the 2SLS results for counties with inelastic and elastic land supply, respectively. Inelastic (elastic) counties have a land unavailability index greater than (less than) 21, following Lutz and Sand (2023). The F statistic reported is the Kleibergen-Paap rk F statistic. County controls include the log of mean income, the log of total population, unemployment rate, the proportion of minority populations, and the percentage of population with post-secondary education. All regressions include county fixed effects and state-year fixed effects. All regressions are weighted by the number of households for each observation. Standard errors in parentheses are clustered at the county level. Significance levels 10%, 5%, 1% are denoted by *, **, and ***, respectively.

Table B9. Income Inequality and House Prices (MSA)

Dependent variable	Gini	log(HPI)			
	First-stage (1)	OLS-all (2)	IV-all (3)	IV-inelastic (4)	IV-elastic (5)
Predicted Gini	0.910***				
	(0.158)				
Gini		-0.505**	7.476***	11.577***	3.963
		(0.251)	(2.298)	(4.357)	(2.934)
Observations	6,454	6,433	6,433	3,228	3,187
F statistic			32.271	18.327	8.948
County controls	Yes	Yes	Yes	Yes	Yes
County fixed effects	Yes	Yes	Yes	Yes	Yes
State-year fixed effects	Yes	Yes	Yes	Yes	Yes

Notes: This table reports the estimation results of the effect of the Gini coefficient on the log of house price index, excluding counties that are not part of metropolitan areas. Column 1 reports the first-stage result. Columns 2 and 3 reports the OLS and 2SLS result for the full sample, respectively. Columns 3 and 4 report the 2SLS results for counties with inelastic and elastic land supply, respectively. Inelastic (elastic) counties have a land unavailability index greater than (less than) 21, following Lutz and Sand (2023). The F statistic reported is the Kleibergen-Paap rk F statistic. County controls include the log of mean income, the log of total population, unemployment rate, the proportion of minority populations, and the percentage of population with post-secondary education. All regressions include county fixed effects and state-year fixed effects. All regressions are weighted by the number of households for each observation. Standard errors in parentheses are clustered at the county level. Significance levels 10%, 5%, 1% are denoted by *, **, and ***, respectively.

Table B10. Income Inequality and Housing Demand

Dependent variable	Gini	log(Origination)		Investment	
	First-stage	OLS	IV	OLS	IV
	(1)	(2)	(3)	(4)	(5)
Predicted Gini	0.926***				
	(0.153)				
Gini		-5.551***	-83.609***	0.323***	6.342***
		(0.571)	(13.978)	(0.082)	(1.215)
Observations	9,277	9,254	9,254	9,264	9,264
F statistic			36.376		36.376
County FE	Yes	Yes	Yes	Yes	Yes
State-year FE	Yes	Yes	Yes	Yes	Yes

Notes: This table reports the estimation results of the effect of the Gini coefficient on housing demand. Column 1 reports the first-stage result. Dependent variables are the log of mortgage originations (columns 2 and 3) and the share of applications for non-owner-occupied properties (columns 4 and 5). The F statistic reported is the Kleibergen-Paap rk F statistic. County controls include the log of mean income, the log of total population, unemployment rate, the proportion of minority populations, and the percentage of population with post-secondary education. All regressions include county fixed effects and state-year fixed effects. All regressions are weighted by the number of households for each observation. Standard errors in parentheses are clustered at the county level. Significance levels 10%, 5%, 1% are denoted by *, **, and ***, respectively.

Table B11. Housing Units and Income Inequality

Dependent variable	ΔLog(Units)	ΔGini	
	First-stage (1)	OLS (2)	IV (3)
Land unavailability	0.0004*** (0.0001)		
ΔLog(Units)		0.034*** (0.013)	0.197* (0.109)
Observations	1,438	1,438	1,438
F statistic			8.771
County controls	Yes	Yes	Yes
State fixed effects	Yes	Yes	Yes

Notes: This table reports the estimation results of the effect of changes in housing stocks on the changes in the Gini coefficient from 1990 to 2017. Column 1 reports the first-stage result, where we use land unavailability as the instrument for the change in housing stocks. Columns 2 and 3 report the OLS and 2SLS result, respectively. The F statistic reported is the Kleibergen-Paap rk F statistic. County controls include the log of mean income, the log of total population, unemployment rate, the proportion of minority populations, and percentage of population with post-secondary education. All regressions include state fixed effects. All regressions are weighted by the number of households for each observation. Robust standard errors reported in parentheses. Significance levels 10%, 5%, 1% are denoted by *, **, and ***, respectively.

Table B12. Income Inequality and California SB35 Votes

Dependent variable	$\mathbb{1}$(Yea)		
	(1)	(2)	(3)
Gini	-7.874**		
	(3.831)		
Top 20% share		-7.760*	
		(4.300)	
Top 5% share			-10.373**
			(4.335)
Observations	37	37	37
Legislative District Controls	Yes	Yes	Yes

Notes: This table reports the estimation results of the relationship between income inequality and the support for California Senate Bill 35. The sample excludes three districts that did not vote. We use three measures of income inequality: the Gini coefficient, the income share of the top 20 percent, and the income share of the top 20 percent. Legislative district controls include the log of mean income, the log of total population, unemployment rate, the proportion of black people, population density, and the share of college graduates. Robust standard errors are reported in parentheses. Significance levels 10%, 5%, 1% are denoted by *, **, and ***, respectively.

Table B13. Income Inequality and Price-Income Ratio

Dependent variable	Price-income ratio			
	OLS-all (1)	IV-all (2)	IV-inelastic (3)	IV-elastic (4)
Gini	1.866* (1.072)	25.490** (10.634)	30.470* (15.956)	19.948** (8.791)
Observations	9,277	9,277	4,654	4,609
F statistic		36.568	20.954	10.250
County controls	Yes	Yes	Yes	Yes
County fixed effects	Yes	Yes	Yes	Yes
State-year fixed effects	Yes	Yes	Yes	Yes

Notes: This table reports the estimation results of the effect of the Gini coefficient on the price-income ratio, which is the ratio between median housing value and median household income. Column 1 reports the first-stage result. Columns 2 reports the 2SLS result for the full sample. Columns 3 and 4 report the 2SLS results for counties with inelastic and elastic land supply, respectively. Inelastic (elastic) counties have a land unavailability index greater than (less than) 21, following Lutz and Sand (2023). The F statistic reported is the Kleibergen-Paap rk F statistic. County controls include the log of mean income, the log of total population, unemployment rate, the proportion of minority populations, and the percentage of population with post-secondary education. All regressions include county fixed effects and state-year fixed effects. All regressions are weighted by the number of households for each observation. Standard errors in parentheses are clustered at the county level. Significance levels 10%, 5%, 1% are denoted by *, **, and ***, respectively.